

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 13:47:14 ; Search time 166.737 Seconds
(without alignments)
274.521 Million cell updates/sec

Title: US-10-019-219A-1

Perfect score: 912

Sequence: 1 TVVRLFLAWLPCMVPCWLP.....WAACGARVKRRFLQLTSLR 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	912	100.0	162	4 AAB31701	Aab31701 Peptide f
2	912	100.0	166	4 AAB31703	Aab31703 Protein e
3	134	14.7	19938	6 ABP76679	Abp76679 Streptomy
4	110.5	12.1	19938	6 ABP76678	Abp76678 Streptomy
5	103.5	11.3	243	4 AAU46489	Aau46489 Propionib
6	103.5	11.3	243	4 ABM43008	Abm43008 Propionib
7	101.5	11.1	270	4 ABG30149	Abg30149 Novel hum
8	100	11.0	306	4 ABG19552	Abg19552 Novel hum
9	100	11.0	19938	6 ABM39838	Abm39838 Streptomy
10	98	10.7	235	6 ADA37505	Ada37505 Human Igg
11	98	10.7	241	2 AAY73982	Aay73982 Human pro
12	98	10.7	470	5 AAU91561	Aau91561 Amino aci
13	97	10.6	157	4 AAU49370	Aau49370 Propionib
14	97	10.6	157	6 ABM45889	Abm45889 Propionib
15	97	10.6	211	6 ABM56588	Abm56588 Propionib
16	97	10.6	332	3 AAY82703	Aay82703 Tick deri
17	96.5	10.6	271	4 AAU58489	Aau58489 Propionib
18	96.5	10.6	271	6 ABM55008	Abm55008 Propionib
19	96.5	10.6	387	2 AAU81589	Aau81589 Protein e
20	96	10.5	408	5 ABG59993	Abg59993 Human DIT
21	95.5	10.5	284	4 ABG05131	Abg05131 Novel hum
22	94	10.3	286	5 ABP69056	Abp69056 Human pol
23	93.5	10.3	180	4 ABG14843	Abg14843 Novel hum
24	93.5	10.3	180	4 ABG12827	Abg12827 Novel hum
25	93.5	10.3	1061	2 AAU87504	Aau87504 Human N-m

26	93.5	10.3	1212	2 AAU87503	Aau87503 Human N-m
27	93.5	10.3	19938	6 ABP76682	Abp76682 Streptomy
28	93	10.2	145	3 AAY87201	Aay87201 Human sec
29	93	10.2	145	4 AAU6178	Aau6178 Human gen
30	93	10.2	145	5 ABG34002	Abg34002 Human sec
31	92.5	10.1	119	4 AAU58360	Aau58360 Propionib
32	92.5	10.1	119	6 ABM54879	Abm54879 Propionib
33	92.5	10.1	331	2 AAU81593	Aau81593 Protein e
34	91.5	10.0	138	4 AAU64233	Aau64233 Propionib
35	91.5	10.0	138	6 ABM60752	Abm60752 Propionib
36	91	10.0	175	4 AAU87255	Aau87255 Novel cen
37	90.5	9.9	108	6 ADB08690	Adb08690 Alloiococ
38	90.5	9.9	243	4 AAU60496	Aau60496 Propionib
39	90.5	9.9	243	6 ABM57015	Abm57015 Propionib
40	90.5	9.9	440	4 ABG05146	Abg05146 Novel hum
41	90	9.9	290	7 ADC86915	Adc86915 Human GPC
42	89.5	9.8	60	2 AAR24223	Aar24223 Partition
43	89.5	9.8	167	4 ABG19252	Abg19252 Novel hum
44	89	9.8	247	4 AAU45904	Aau45904 Propionib
45	89	9.8	247	6 ABM43423	Abm43423 Propionib

ALIGNMENTS

RESULT 1

ID AAB31701 standard; peptide; 162 AA.

XX AC AAB31701;

XX 30-APR-2001 (first entry)

XX Peptide fragment of a human intestinal carboxylesterase (iCE).

XX Intestinal carboxylesterase; iCE; tumour; cytotoxic T lymphocyte; CTL;

XX cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;

XX tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.

XX Homo sapiens.

XX WO200100784-A2.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-FR001791.

XX 28-JUN-1999; 99FR-00008224.

XX (INSR) INST ROUSSY GUSTAVE.

XX Ronsin C, Scott V, Triebel F;

XX WPI; 2001-112443/12.

XX New peptides and its encoding nucleic acid derived from intestinal carboxylesterase, useful as immunostimulants for treating cancer.

XX Claim 1; Page 3; 53pp; French.

XX The present sequence is derived from a human intestinal carboxylesterase (iCE) polypeptide. iCE induces specific cytotoxic T lymphocytes (CTL) and secretion by these CTL of cytotoxic factors, e.g. interleukin-2, interferon gamma and tumour necrosis factor. iCE polypeptides and polynucleotides are used for treating cancer, by in vivo or in vitro immunisation, particularly solid cancers and most especially hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used to stimulate the immune system, and to increase, in culture, the production of associated-associated CTL, for reinjection, and/or to induce secretion of cytotoxic factors from CTL. Dendritic cells loaded with iCE are used to induce such CTL in cultures

XX Sequence 162 AA;

Query Match 100.0%; Score 912; DB 4; Length 162;
 Best Local Similarity 100.0%; Pred. No. 1.3e-78;
 Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVVRLFLAWLPCMMVPCWLPWRTWWSSSTAWVSWASSALETSTQPATGATWTKWLHYA 60
 DB 1 TVVRLFLAWLPCMMVPCWLPWRTWWSSSTAWVSWASSALETSTQPATGATWTKWLHYA 60
 QY 61 GSSRISPTLEATLTVPFLASLRVARVCLRLCCPPYKDSSTEPSWRVAVPSCPASLPAQ 120
 DB 61 GSSRISPTLEATLTVPFLASLRVARVCLRLCCPPYKDSSTEPSWRVAVPSCPASLPAQ 120
 QY 121 LMSPPRWPTCLPVTKLTLPWMAACGARVKRRFLQLTSLR 162
 DB 121 LMSPPRWPTCLPVTKLTLPWMAACGARVKRRFLQLTSLR 162

RESULT 2

AAB31703
 ID AAB31703 standard; protein; 166 AA.

AC AAB31703;

DT 30-APR-2001 (first entry)

XX Protein encoded by an intestinal carboxylesterase (iCE) cDNA.

XX Intestinal carboxylesterase; iCE; tumour; cytotoxic T lymphocyte; CTL;
 KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
 KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.

XX Homo sapiens.

XX WO200100784-A2.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-FR001791.

XX 28-JUN-1999; 99FR-00008224.

XX (INSR) INST ROUSSY GUSTAVE.

XX Ronsin C, Scott V, Triebel F;

XX WPI; 2001-112443/12.

XX N-PSDB; AAP25258.

XX New peptides and its encoding nucleic acid derived from intestinal
 PT carboxylesterase, useful as immunostimulants for treating cancer.

XX Disclosure; Fig 8A; 53pp; French.

XX The present sequence is encoded by the coding region of human intestinal
 CC carboxylesterase (iCE) gene. iCE induces specific cytotoxic T
 CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors, e.g.
 CC interleukin-2, interferon gamma and tumour necrosis factor. iCE
 CC polypeptides and polynucleotides are used for treating cancer, by in vivo
 CC or in vitro immunisation, particularly solid cancers and most especially
 CC hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used
 CC to stimulate the immune system, and to increase, in culture, the
 CC production of associated-associated CTL, for reinjection, and/or to
 CC induce secretion of cytotoxic factors from CTL. Dendritic cells loaded
 CC with iCE are used to induce such CTL in cultures

XX Sequence 166 AA;

XX Query Match 100.0%; Score 912; DB 4; Length 166;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-78;

XX Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVVRLFLAWLPCMMVPCWLPWRTWWSSSTAWVSWASSALETSTQPATGATWTKWLHYA 60

DB 5 TVVRLFLAWLPCMMVPCWLPWRTWWSSSTAWVSWASSALETSTQPATGATWTKWLHYA 64
 QY 61 GSSRISPTLEATLTVPFLASLRVARVCLRLCCPPYKDSSTEPSWRVAVPSCPASLPAQ 120
 DB 65 GSSRISPTLEATLTVPFLASLRVARVCLRLCCPPYKDSSTEPSWRVAVPSCPASLPAQ 124
 QY 121 LMSPPRWPTCLPVTKLTLPWMAACGARVKRRFLQLTSLR 162
 DB 125 LMSPPRWPTCLPVTKLTLPWMAACGARVKRRFLQLTSLR 166

RESULT 3

ABP76679

ID ABP76679 standard; protein; 19938 AA.

XX AC ABP76679;

XX 26-FEB-2003 (first entry)

XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 3.

XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
 KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.

XX Streptomyces viridochromogenes.

XX WO200268436-A1.

XX 06-SEP-2002.

XX 24-AUG-2001; 2001WO-EP009815.

XX 25-FEB-2001; 2001DE-01009166.

XX (COMB-) COMBINATURE BIOPHARM AG.

XX Weitnauer G, Muehlenweg A, Trefzer A, Bechthold A;

XX WPI; 2003-018650/01.

XX N-PSDB; ABZ37515.

XX New avilamycin derivatives, useful for treatment of infections, and
 PT nucleic acid encoding avilamycin synthesis enzymes.

XX Example 1; Page 68-301; 319pp; German.

XX The invention relates to avilamycin derivatives (I) with antibacterial,
 CC virucide, protozoacide and fungicide activity. (I) are useful for
 CC treatment of infections (bacterial, viral, protozoal or fungal), in human
 CC or veterinary medicine, particularly where caused by Staphylococcus
 CC aureus. (I) are more hydrophilic than known avilamycins. The present
 CC sequence is that of an avilamycin synthesis enzyme from the Streptomyces
 CC viridochromogenes Avilamycin A biosynthetic gene cluster (ABZ37515-
 CC ABZ37516)

XX Sequence 19938 AA;

XX Query Match

XX Best Local Similarity 14.7%; Score 134; DB 6; Length 19938;

XX Matches 47; Conservative 15; Mismatches 68; Indels 64; Gaps 7;

QY 20 PARTWW-----WSSSSTAWVSWASSALET---STOPATGATWTKWLHYAGSS--- 63

DB 4681 PARTCAARSPSPAPWSAPATWNTSAATTACWRPTPTTGPANSEWTFPPSSPRPT 4740

QY 64 -----RISPLEATLTVPFLASLRVARVCL-----RLLC 93

DB 4741 RRASXSPTSSHTTCSAGAAPRSSPRSCSTTCRVWSSCGRSAAACXRTTASGXPSRAIC 4800

QY 94 P-----PYPKDSSTEPSWRV-----AWPSCPASL-----PAQLMSSPRWPTCLPVT 135

DB 4801 PRCSTPAPTWSATSTWTTTGCARSGWSPAPAXRWMPSPXPPSTAEASRCWPAAAPLA 4860


```

Query Match      11.0%; Score 100; DB 6; Length 19938;
Best Local Similarity 22.2%; Pred. No. 38;
Matches 46; Conservative 14; Mismatches 43; Indels 104; Gaps 10;

QY 9 WLPQWVPCWLP-----WRTMWSSSTAWVS-----WA 37
DQ 11462 WWPASMXPARPCAGTSASPASTRSRSPAPARTWRCGAATWAGRSVAVKNCWS 11521
QY 38 SSALETSTQPATGATWTKWHLHYAGSSRISPTLEATLTVPFLASLRVARVCLRLCPLYP 97
DQ 11522 SST--XPREPTSGSTTRAVCGAGST-----WP 11547
QY 98 KDSSTEPQWVWAPVS-CPASLPAQLMSSPRW---W-----PVCLPVTXLTLR-----140
DQ 11548 PDSFPCPCS-SWTSPPPASTRATRSGRVSRVSRAPPCSP--RSTWRRPTGSPT 11604
QY 141 -----PWWAACA 148
DQ 11605 SRXSTRAAGSRAPRPXPSPWAAATGS 11631

RESULT 10
ADA37505
ID ADA37505 standard; protein; 235 AA.
XX
AC ADA37505;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human Igg fragment crystalline region ORF 2.
XX
KW Cytostatic; immunosuppressive; antiallergic; antimicrobial;
KW immune system; FcRn; heavy chain constant region; receptor;
KW autoimmune disease; tumour; epithelial cells; immunoglobulin; Igg Fc;
KW human; allergy; fragment crystalline.
XX
OS Homo sapiens.
XX
FH Key
FT Misc-difference 34
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT Misc-difference 48
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT Misc-difference 96
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT Misc-difference 140
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT Misc-difference 147
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT Misc-difference 209
FT /label= UNKNOWN
FT /note= "Encoded by in frame stop codon TGA"
FT
FT US2002192222-A1.
XX
XX 19-DEC-2002.
XX
XX 08-AUG-2002; 2002US-00215297.
XX
XX 17-JAN-1995; 95US-00374159.
XX 29-DEC-1995; 95US-00578171.
XX 24-JUL-1997; 97US-00899856.
XX 24-JUL-1998; 98US-00122144.
XX
XX (BLUM/) BLUMBERG R S.
XX (SIMI/) SIMISTER N E.
XX (LENC/) LENCER W I.

```

```

XX PI Blumberg RS, Simister NE, Lencer WI;
XX WPI; 2003-657224/62.
XX N-PSDB; ADA37501.
XX
XX Modulating immune system of a mammal, by administering to epithelial
XX barrier of a mammal, a conjugate of FcRn binding partner and an antigen
XX e.g., antigen of a pathogen, autoimmune disease, allergen or tumor.
XX
XX Disclosure; Fig 1; 25pp; English.
XX
XX The invention relates to modulating (M1) the immune system of a mammal,
XX comprising administering to an epithelial barrier of a mammal in need of
XX such immune modulation, a conjugate of an antigen and an FcRn (Fragment
XX crystalline receptor, the immunoglobulin receptor recognising the heavy
XX chain constant region) binding partner, where the antigen is an antigen
XX of a pathogen, autoimmune disease, allergen or tumour. Also included are
XX a pharmaceutical preparation (comprising, a conjugate of an antigen and
XX FcRn binding partner, where the antigen is characteristic of a tumour,
XX and a carrier, where the conjugate is present in an amount effective for
XX modulating the immune response of a mammal) orally delivering molecules
XX to a mammal (involving administering to the mammal, a conjugate of
XX therapeutic and a FcRn binding partner targeted to epithelial cells
XX expressing FcRn, and delivering (M2) molecules to a mammal, (involving
XX administering to the mammal, a conjugate of a bioactive substance and a
XX FcRn binding partner targeted to epithelial cells expressing a FcRn).
XX (M1) is useful for modulating the immune system of a mammal. (M2) is
XX useful for delivering bioactive substances such as cells, viruses,
XX vectors, proteins, peptides, nucleic acids, polysaccharides and
XX organic and inorganic drugs exerting a biological effect when
XX administered to a mammal. (M2) is useful for delivering drugs such as
XX antineoplastic compounds, immunoactive compounds, antimicrobial agents,
XX parasiticides, haematologic compounds, cardiovascular drugs, respiratory
XX drugs, neuromuscular blocking drugs, etc. to epithelial borders. The
XX method permits more effective strategies for immunising humans. The
XX present sequence is the protein encoded by open reading frame 1 (ORF2) of
XX the DNA appearing as ADA37501 stated to encode human Igg Fc which is
XX recognised by FcRn i.e. is an FcRn binding partner.
XX
XX Sequence 235 AA;
XX
XX Query Match      10.7%; Score 98; DB 6; Length 235;
XX Best Local Similarity 25.8%; Pred. No. 0.32;
XX Matches 42; Conservative 14; Mismatches 65; Indels 42; Gaps 9;
XX
XX QY 24 WWW-----SSSTAWVSWASSALETSTQPATGATWTKWHLHYAGS-SRISPTLEATLT 74
XX DQ 44 WWWTXATKTLRSSTGTWTAWRCIMPRQSRGRSSTTARTVWSASSPSCTRTGMAKSTSA 103
XX QY 75 VSPFLASLRVARV-----CLRLCPL-----PYPKDSSTSPSWRVAMPSPASL 117
XX DQ 104 RSFTKPSQPPSRKPKGSPKXGSPENHRCPT--CFHFGMSXPRTSAXPAMSKA--SIPATS 159
XX QY 118 P--AQLMSS-----PRWPTCLPVTXLTLRPWWAACGAR 149
XX DQ 160 PWSGRAVGSRRITTRPRLPCWTPTA-PSSSTASSPWPTRAGGR 201
XX
XX RESULT 11
XX AAY73982
XX ID AAY73982 standard; protein; 241 AA.
XX
XX AC AAY73982;
XX
XX 14-MAR-2000 (first entry)
XX
XX Human prostate tumor EST fragment derived protein #169.
XX
XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX treatment.
XX
XX

```


XX PA (CORI-) CORIXA CORP.
 XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 XX PI Barth B, Vallie-Douglass J;
 XX DR WPI: 2003-381789/36.
 XX DR N-PSDB; ACF64474.
 XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 XX PT or for stimulating an immune response specific for a P. acnes protein.
 XX PS Example 1; SEQ ID NO 10565; 1481pp; English.
 XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 XX CC encoding a Propionibacterium acnes protein. The invention also relates to
 XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 XX CC immunogenic fragments of P. acnes polypeptides. The invention
 XX CC additionally encompasses expression vectors and host cells comprising a
 XX CC polynucleotide of the invention; antibodies against polypeptides of the
 XX CC invention; fusion proteins comprising a polypeptide of the invention; a
 XX CC method for stimulating an immune response specific for a P. acnes
 XX CC polypeptide and an isolated T cell population comprising T cells prepared
 XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 XX CC antigen-presenting cells that express the polypeptide); a method and kit
 XX CC for detecting or determining the presence or absence of P. acnes in a
 XX CC patient; and a method for inhibiting the development of P. acnes in a
 XX CC polypeptides are useful for diagnosing, preventing or treating acne
 XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
 XX CC protein. The polynucleotides can also be used as probes or primers for
 XX CC nucleic acid hybridisation. The vaccine composition is useful for the
 XX CC stimulation of an immune response against P. acnes, or for treating acne,
 XX CC and the kit is useful for performing a diagnostic assay. The present
 XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 XX CC reading frame) contained within the P. acnes polynucleotides of the
 XX CC invention. Note: The sequence data for this patent did not form part of
 XX CC the printed specification, but was obtained in electronic format directly
 XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX CC
 XX CC Sequence 157 AA;
 Query Match 10.6%; Score 97; DB 6; Length 157;
 Best Local Similarity 28.9%; Pred. No. 0.25; Indels 36; Gaps 9;
 Matches 39; Conservative 17; Mismatches 43; Indels 36; Gaps 9;
 QY 26 WSSSSTAW---VSWASSALETSTQPATGATWTKLHYAGSSRISPT---LEATLTVSPFL 79
 Db 1 WRICAPAWGASVVRWARTTSATCA-PSTASISWR-----SRMPPTARLSPVTPFPSP 52
 QY 80 ASLRVARVCLRLCPYPKDSSTE---PSW-RVAMPSCFASLPALQMSRPRWPTCLPV 134
 Db 53 TSMRSSRL-VSTTVSSLPRRSSTRSSVSWLRGSGTP-----SSRSRWPTTLRL 100
 QY 135 TKLTLRP-----WVA 144
 Db 101 RLVLRLPLRARVWVA 115
 RESULT 15
 ABM65658
 ID ABM65658 standard; protein; 211 AA.
 XX AC
 XX AC ABM65658;
 XX XX
 XX XX 20-OCT-2003 (first entry)
 XX DE Propionibacterium acnes immunogenic polypeptide #10334.
 XX XX

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI: 2001-616774/71.
 XX DR N-PSDB; AAS59545.
 XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 XX PT vaccinating against and diagnosing infections, especially useful for
 XX PT treating acne vulgaris.
 XX PS Example 1; SEQ ID NO 10565; 1069pp; English.
 XX CC Sequences AU9105-AU68017 represent Propionibacterium acnes immunogenic
 XX CC polypeptides. The proteins and their associated DNA sequences are used in
 XX CC the treatment, prevention and diagnosis of medical conditions caused by
 XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 XX CC P. acnes is also involved in infections of bone, joints and the central
 XX CC nervous system, however it is particularly involved in the inflammatory
 XX CC lesions associated with acne vulgaris. A method for detecting the
 XX CC presence or absence of P. acnes in a patient comprises contacting a
 XX CC sample with a binding agent that binds to the proteins of the invention
 XX CC and determining the amount of bound protein in the sample. The
 XX CC polypeptides may be used as antigens in the production of antibodies
 XX CC specific for P. acnes proteins. These antibodies can be used to
 XX CC downregulate expression and activity of P. acnes polypeptides and
 XX CC therefore treat P. acnes infections. The antibodies may also be used as
 XX CC diagnostic agents for determining P. acnes presence, for example, by
 XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 XX CC this patent did not form part of the printed specification, but was
 XX CC obtained in electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences
 XX CC
 XX CC Sequence 157 AA;
 Query Match 10.6%; Score 97; DB 4; Length 157;
 Best Local Similarity 28.9%; Pred. No. 0.25; Indels 36; Gaps 9;
 Matches 39; Conservative 17; Mismatches 43; Indels 36; Gaps 9;
 QY 26 WSSSSTAW---VSWASSALETSTQPATGATWTKLHYAGSSRISPT---LEATLTVSPFL 79
 Db 1 WRICAPAWGASVVRWARTTSATCA-PSTASISWR-----SRMPPTARLSPVTPFPSP 52
 QY 80 ASLRVARVCLRLCPYPKDSSTE---PSW-RVAMPSCFASLPALQMSRPRWPTCLPV 134
 Db 53 TSMRSSRL-VSTTVSSLPRRSSTRSSVSWLRGSGTP-----SSRSRWPTTLRL 100
 QY 135 TKLTLRP-----WVA 144
 Db 101 RLVLRLPLRARVWVA 115
 RESULT 14
 ABM45889
 ID ABM45889 standard; protein; 157 AA.
 XX AC
 XX AC ABM45889;
 XX XX
 XX XX 20-OCT-2003 (first entry)
 XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #10565.
 XX XX
 XX XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 XX KW immunostimulant; immune response; vaccine.
 XX XX
 XX XX Propionibacterium acnes.
 XX OS
 XX OS WO2003033515-A1.
 XX FN
 XX XX
 XX XX 24-APR-2003.
 XX PD
 XX XX 11-OCT-2002; 2002WO-US032727.
 XX PF
 XX XX 15-OCT-2001; 2001US-00978825.
 XX FR

KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine; immunogenic.
OS Propionibacterium acnes.

FN WO2003033515-A1.

PD 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.

PS Claim 7; SEQ ID NO 30334; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed P. acnes polypeptide which is
CC thought to contain an immunogenic region. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 211 AA;

Query Match 10.6%; Score 97; DB 6; Length 211;
Best Local Similarity 28.9%; Pred. No. 0.35;
Matches 39; Conservative 17; Mismatches 43; Indels 36; Gaps 9;
QY 26 WSSSSTAW---VSWASSALETSTQATGATWTKMLHYAGSSRISPT---LEATLTVSPFL 79
Db 1 WRICAPAWASVWARTSATCA-PTASINSR-----SRMPPARTRLESPTVFPSP 52
QY 80 ASLRVARVRLCLCPYPKDSSTE---PSW--RVAMPSCPASLPALQIMSSPRWWPTCLPV 134
Db 53 TSMRSSRL-VSTTVSSLPRSSRRSSVSLRGSTP-----SSRSRWPTLRL 100
QY 135 TKLTLRP-----WVA 144
Db 101 RLWMLRPLRLRVWVA 115

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:16:17 ; Search time 42.6316 Seconds
(without alignments)
196.179 Million cell updates/sec

Title: US-10-019-219A-1

Perfect score: 912

Sequence: 1 TVVRLFLAWLPCMVPCWLP.....WAACGARVKRFLQLTSLR 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	12.5	169	4	US-09-252-991A-32019
2	114	12.5	299	4	US-09-252-991A-17588
3	113.5	12.4	341	4	US-09-252-991A-32424
4	113	12.4	243	4	US-09-252-991A-19853
5	108	11.8	247	4	US-09-252-991A-26899
6	104.5	11.5	191	4	US-09-252-991A-25365
7	104.5	11.5	222	4	US-09-252-991A-32998
8	103.5	11.3	431	4	US-09-252-991A-25721
9	103	11.3	246	4	US-09-252-991A-24028
10	101	11.1	208	4	US-09-252-991A-32166
11	100.5	11.0	160	4	US-09-252-991A-30765
12	100.5	11.0	178	4	US-09-252-991A-31386
13	100.5	11.0	297	4	US-09-252-991A-28842
14	100	11.0	1650	4	US-09-252-991A-21798
15	99.5	10.9	166	4	US-09-252-991A-24322
16	99.5	10.9	302	4	US-09-252-991A-19529
17	98.5	10.8	171	4	US-09-252-991A-16782
18	98.5	10.8	171	4	US-09-252-991A-25789
19	98.5	10.8	184	4	US-09-252-991A-28187
20	98	10.7	299	4	US-09-252-991A-17878
21	98	10.7	303	4	US-09-252-991A-22757
22	98	10.7	312	4	US-09-252-991A-32997
23	97	10.6	214	4	US-09-252-991A-32533
24	97	10.6	394	4	US-09-252-991A-27774
25	97	10.6	701	4	US-09-252-991A-32462
26	96.5	10.6	270	4	US-09-252-991A-19031
27	96.5	10.6	327	4	US-09-252-991A-30618

```

28 96.5 10.6 405 4 US-09-252-991A-30240 Sequence 30240, A
29 95.5 10.5 173 4 US-09-252-991A-32359 Sequence 32359, A
30 95 10.4 185 4 US-09-252-991A-31599 Sequence 31599, A
31 95 10.4 194 4 US-09-252-991A-31583 Sequence 31583, A
32 94.5 10.4 134 4 US-09-252-991A-25639 Sequence 25639, A
33 94.5 10.4 194 4 US-09-252-991A-24154 Sequence 24154, A
34 94.5 10.4 239 4 US-09-252-991A-30037 Sequence 30037, A
35 94.5 10.4 426 4 US-09-252-991A-18298 Sequence 18298, A
36 94 10.3 188 4 US-09-252-991A-32934 Sequence 32934, A
37 94 10.3 203 4 US-09-252-991A-21364 Sequence 21364, A
38 94 10.3 239 4 US-09-252-991A-19211 Sequence 19211, A
39 93.5 10.3 109 4 US-09-252-991A-31463 Sequence 31463, A
40 93.5 10.3 369 4 US-09-252-991A-20790 Sequence 20790, A
41 93.5 10.3 1093 4 US-09-252-991A-21827 Sequence 21827, A
42 93 10.2 145 4 US-09-482-273-240 Sequence 240, App
43 93 10.2 225 4 US-09-252-991A-30156 Sequence 30156, A
44 93 10.2 235 4 US-09-252-991A-18300 Sequence 18300, A
45 92.5 10.1 169 4 US-09-252-991A-30563 Sequence 30563, A

```

ALIGNMENTS

```

RESULT 1
US-09-252-991A-32019
; Sequence 32019, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32019
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32019

```

```

Query Match 12.5%; Score 114; DB 4; Length 169;
Best Local Similarity 27.0%; Pred. No. 0.00035;
Matches 44; Conservative 14; Mismatches 67; Indels 38; Gaps 9;

QY 21 WRTWWSSSTANVSWASSALETSTQATGAT----WTK-WLHYAGSSRISP-----67
Db 1 WKNW----PATATRSTPSSATATRRPACPTARCCPWTGCM---SNTCVGPPTACRRP 52
QY 68 ---TLEATLTVSPFLASLVARVCLLLCPYPKDSSTEPSWVAMPSCASLP-----119
Db 53 CARAMSATISTASTASCTPPST---CPPRPTAPST-ARRYGWPTACSSMTACRRPAR 106
QY 120 QLMSSPRWPTCLPVTKLTLRPWWAACGARVKRRFLQLTSLR 162
Db 107 SRTGSPWTFP---PATSSTPARWACPSAARPPARSASWSTDAAR 146

```

```

RESULT 2
US-09-252-991A-17588
; Sequence 17588, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18

```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17588
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17588

Query Match 12.5%; Score 114; DB 4; Length 299;
Best Local Similarity 28.7%; Pred. No. 0.00071;
Matches 43; Conservative 13; Mismatches 62; Indels 32; Gaps 7;

QY 10 LPCMMVPCWLPWRTW---LPWRTWSSSTAWVSWASSALETSTQPATGATWTKWLHYAGSS 63
DB 6 LPCWRFCTCWPRSPWPPAADCPTSAWPRPWSAATSPPRATTREMPRWVPTPRS 65
QY 64 RISPTLEATLTVSPFLASLRVARVCLRLCPYPKDSSTEPSWRVWVWPCFASLPAQLMS 123
DB 66 -ASPTCRSPAATAP-----SRPRKST-----ARPRPRRGPARALA 102
QY 124 SPRWPTCLPVTKLTLRPWMAACGARVKRR 153
DB 103 S-GMRPSPAPATPSTARP--SACSARPSNR 129

RESULT 3

US-09-252-991A-32424
; Sequence 32424, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32424
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32424

Query Match 12.4%; Score 113.5; DB 4; Length 341;
Best Local Similarity 25.1%; Pred. No. 0.00095;
Matches 50; Conservative 15; Mismatches 77; Indels 57; Gaps 9;

QY 9 WPCMMVPCWLP-----W-----RTWSSSSTAWVSWASSALETSS 44
DB 66 WRPAAPAPWIPFYCLPRSPAKRGSAWCRPPRSIRPTWPGSNCRCTGSAIVAPAGT 125
QY 45 TOPATCAT---WTKWLHYAGSSRISPTLEATL---TVSPFLASLRVARVCLRLCPYP 96
DB 126 SSFRCATPTSATWRCFMRSSATRAPSPGWCCTTCGTASPAHHSWSTARAA---ATPTF 182
QY 97 P-----KDSST-BPSWRVWVWPCFASLPAQLMSSPRWPT-CLPVTKLTLR 140
DB 183 PGCTRTSTTTANGSRSDRSTCRPSARRGFFWCRPAPPRAGATSPPRWPTWCSRRPRRL 242
QY 141 PW-----WAACGARVKRR 153
DB 243 PMNCAATFFAAPTATVARR 261

RESULT 4

US-09-252-991A-19853

; Sequence 19853, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 19853

; LENGTH: 243

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19853

Query Match 12.4%; Score 113; DB 4; Length 243;

Best Local Similarity 27.0%; Pred. No. 0.0007;
Matches 47; Conservative 16; Mismatches 63; Indels 48; Gaps 11;

QY 10 LPCMMVPCWLPWRTW-----WSSSSTAWVSWASSALETSTQPATGATW-----TKWLHY 59
DB 11 MPATEAACAPWR-WPPSMWAPASAPIRASATWASGVGTS--PRRPSSWSTTCTSTSLAHC 67
QY 60 AGS-----SRISPTLEATL---TVSPFLASLRVA-----RVCLRLCPYPKDSSTE 103
DB 68 ANSPGCRACRCNTSSPTARSTCTWGTSTRSPRACSSRPCNWNRCNCIACPVRRPGASA- 126
QY 104 PSWRVWVWPCFASLPAQLMSSPRWPT-----TCLPVTKLTLR 141
DB 127 GNASRW--CASSTPTATTAAAPSSPGSPSRSRSPRSRSCAPARARCAP 178

RESULT 5

US-09-252-991A-26899

; Sequence 26899, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26899

; LENGTH: 247

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26899

Query Match 11.8%; Score 108; DB 4; Length 247;

Best Local Similarity 25.3%; Pred. No. 0.0023;
Matches 45; Conservative 11; Mismatches 62; Indels 60; Gaps 6;

QY 11 PCMMVPC-----WLPWRTWSSSSTAWVSWASSALETSTQPATGATWTKWLHYA 60
DB 71 PWSAGCSASRACNCKSWIPASPPRCSPSASAWPSPCARCNRWPPRCSPWPTGPTA 130
QY 61 GSSRISPTLEATLTVSPFLASLRVARVCLRLCPYPKDSSTEP-----SWR----- 107
DB 131 WSCRTSPAPGCTRSVS-----SPPPPR--CTPPATWCAWRRRRPRPW 172
QY 108 -----VANWPCFASLPAQLMSSPRWPTCLPVTKLTLRPWMAACGARVKRR 153

Db 173 VSPGWPRRAGSGCWPAEPPLPQRLISTRRPWCT-----RRPWPVCASIRASGR 223

RESULT 6

US-09-252-991A-25365

; Sequence 25365, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25365

; LENGTH: 191

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25365

Query Match 11.5%; Score 104.5; DB 4; Length 191;

Best Local Similarity 25.1%; Pred. No. 0.0038;

Matches 44; Conservative 13; Mismatches 61; Indels 57; Gaps 7;

QY 8 AWLPCMVMPWLPWRTWWSSSTAWVSWASSALETSTQPATGATWTKWLHYAGSSRISP 67

Db 4 ATIPANCAPCTANCRCWCWITAS---WPSWI-----PTWSAWITPTSNRPST 47

QY 68 T-----LEATLVSPFLASLRV-----ARVCLRLCLPPYPKDSSTE 103

Db 48 TCSTAASATSCWSPNRSTAPARGSGSRPSPMPGAGDRSRGRCK---PTTSSATACA 103

QY 104 PSWRVAMP-----SCPASLPAQLMSSP---RWPTCLPVTCLTLRPWAA 145

Db 104 PSFPAAPGRRRCSPATASPSPAPPANCATSAAGCSMRSCWPNWTNSTGTHWAA 158

RESULT 7

US-09-252-991A-32998

; Sequence 32998, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 32998

; LENGTH: 222

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32998

Query Match 11.5%; Score 104.5; DB 4; Length 222;

Best Local Similarity 24.7%; Pred. No. 0.0046;

Matches 47; Conservative 11; Mismatches 71; Indels 61; Gaps 8;

QY 12 CMVVP--CWLPPWTTWWSSSTAWVSWAS-----SALETSTQPATGATWTKWLHYAGSSR 64

Db 37 CRAIPRCCSANWARVPWPTGYLTSTAPWTSRPPCRPACPTCCSTAPPASRWAW----- 88

QY 65 ISPTLEAT-----LTVSPFLASLRVARVCLRLCLPPYPKDS-----STEP- 104

Db 89 -PPTCRRRTCKGSRRRPACACSTSRARRSPNCANTCAARTSPFKPRSSPRAPTCSRSTPA 147

QY 105 SWRVAMPSCPASLPAQLMSSP-----RWPTCLPVTCLTLRP 141

Db 148 AVRCACARCTASRTATSSSTCTCTPCVPCWNRSPGRCRPSCHWNFTC--ATSTRTT 205

QY 142 WWAACGARVK 151

Db 206 RPASSSSRRV 215

RESULT 8

US-09-252-991A-25721

; Sequence 25721, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25721

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25721

Query Match 11.3%; Score 103.5; DB 4; Length 431;

Best Local Similarity 24.3%; Pred. No. 0.013;

Matches 42; Conservative 12; Mismatches 70; Indels 49; Gaps 8;

QY 8 AWLP--C-MVPCW-----LPWRTWWSSSTAWVSWASSALETSTQPA 48

Db 251 AWPTACRRPPPCWPGPRSPGSRATDSLWKWRDWSRGACSTSPSWFDR-----S 302

QY 49 TGATWTKWLHYAGSSRSISPTLEATLVSPFLASLRVARVCLRLCLPPYPKDSSTEPSWRV 108

Db 303 PGRRCARWRMSARAASASGSTPSRRSSP-----RRQRPCAIPLPPTGCRSPSWASPGWRA 358

QY 109 AWPSCPASLPAQLMSSPRW-----WPTCLPVTCLTLRPWAAAC 146

Db 359 NWMA--PGWFSWIAVRRRWAGSSSCRCCTAKAGWPACGSAATWTER-WPGAC 408

RESULT 9

US-09-252-991A-24028

; Sequence 24028, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 24028

; LENGTH: 246

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24028

```

Query Match      11.3%; Score 103; DB 4; Length 246;
Best Local Similarity 26.4%; Pred. No. 0.0074;
Matches 39; Conservative 20; Mismatches 63; Indels 26; Gaps 7;

QY      17  CMLPWTWWSSTSTA-----WVSWASSALETSTOPAT-GATWTWKWLHYAGSSRI 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      79  CWPSSSTPRPSTGATTTVPGRNSCRWSNW-NSAWRSPSPSTRKACWAT-----GSAARP 133

QY      66  SPTLEATLTVSPFLASLRVARVCLRLICPPYPKDDSTEPSRWRVWVWPCPASLPAQLMSSP 125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      134  RPTSTASIAIRTSAAAGVFTACSWRRPVRSGSAATCSAKARSATCSCSAAP-RSTTSP 192

QY      126  RW-----WPTCLPVTKLTLRPWAAACGAR 149
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      193  RWPDSVAMPFTA-----RRLSETWLRCCAR 215

RESULT 10
US-09-252-991A-32166
; Sequence 32166, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32166
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32166

Query Match      11.1%; Score 101; DB 4; Length 208;
Best Local Similarity 29.08; Pred. No. 0.0096;
Matches 42; Conservative 18; Mismatches 57; Indels 28; Gaps 7;

QY      21  WRTWWSSTSTAWVSWASSALETSTOPATGATWTWKWLHYAGSSRISPTLEATLTVSPFLA 80
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      70  WRTCRSPASVTRSTSNITSA--STTMAACTASRSTWPAIAAIRRTSPARSST---AP-C 122

QY      81  SLRVARVCLRLICPPYPKDDSTEPSR-WVWPCPASLPAQLMSSPRWPTCLPVTKLTL 139
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      123  STRTT-----PTSGATPSTATVARPTPTPTTAVSAGPREWSP-----SR 164

QY      140  RPW--WAACGARVKRRFLQLTSLSR 162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      165  RSWTPPAAWARIKWCASITITAR 189

RESULT 11
US-09-252-991A-30765
; Sequence 30765, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30765

```

```

: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 28842
: LENGTH: 297
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28842

```

Query Match	11.0%	Score 100.5	DB 4	Length 297
Best Local Similarity	26.3%	Pred. No. 0.017		
Matches	47	Conservative 11	Mismatches 44	Indels 77
				Gaps 14
QY	12	CMNVPCWLPWRTWWSGST	---AWYSW----	ASSALETSTQPATGATWTKLHYAG 61
Db	49	CCGAPC-----	WWASSATSGNAANSWWRSTASCASTNSPAPIT-----	89
QY	62	SKRISPTLEATLVSPFLASLRVARV-----	CLRLCPVPYKDSSTEPSRVRVAPS-----	112
Db	90	-SCASPTI-AWNSGSP--	ATTRACAVRCSVCDNCCASCAMRRHRSWRKPCWR-CWPST	144
QY	113	-----CPASLPAQLMSSP-----	---RWNPT---CLPVTKLTFLREW-----	146
Db	145	VRCASPDSTGCATAPRS--	RNPWRRSARCRRTPTICWCAPA-----WSAANSAC	194

```

RESULT 14
US-09-252-991A-21798
; Sequence 21798, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21798
; LENGTH: 1650
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21798

```

```

Query Match      11.0%; Score 100; DB 4; Length 1650;
Best Local Similarity 26.5%; Pred. No. 0.17;
Matches 54; Conservative 14; Mismatches 48; Indels 88; Gaps 15;

16 PCWL-----PWRT-----WWSSSSTAMVSWASSALE--TSTQATGATWTK-----WLHYA 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
87 PFWLSAPPIRSELSPWMS-----VPWRLSLPACTSRAPA-----WLKIRSCWLRRA 136
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

61 GSSR-----ISPTLEATLVSPFFIASL--RVARVCLR--LLCPFY----- 96
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

137 PPRRRRLASWILARVEALLSRLPVTIVSSSPWLTSLPRLSTLALSSRSPPLARRAAPA 196
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

97 -----PKDSSTPSNRVAVPSC-----PASLPAQLMS----- 124
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

197 LTRAPRWASSRP--WLINWPSLLSRLIAPLSRLLPPIRPPALFSSVSSRSKLSARCAETR 255
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

125 PRW-----WPTCLPVTXLT--LRPW 142
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

256 PPWPISSFFPAMLRVTSLSPTMPW 279
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 15
US-09-252-991A-24322
; Sequence 24322, Application US/09252991A

```

; Patent No. 6551795
;
; GENERAL INFORMATION:
;
; APPLICANT: Marc J. Rubenfield et al.
;
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;
; FILE REFERENCE: 107196.136
;
; CURRENT APPLICATION NUMBER: US/09/252,991A
;
; PRIORITY FILING DATE: 1999-02-18
;
; PRIOR APPLICATION NUMBER: US 60/074,788
;
; PRIORITY FILING DATE: 1998-02-18
;
; PRIOR APPLICATION NUMBER: US 60/094,190
;
; PRIORITY FILING DATE: 1998-07-27
;
; NUMBER OF SEQ ID NOS: 33142
;
; SEQ ID NO 24322
;
; LENGTH: 166
;
; TYPE: PRT
;
; ORGANISM: Pseudomonas aeruginosa
;
; US-09-252-991A-24322

```

Query Match	10.9%	Score 99.5;	DB 4;	Length 166;
Best Local Similarity	25.0%;	Pred.No.	0.01;	
Matches	42;	Conservative	14;	Mismatches 53; Indels 59; Gaps 9;
QY	25	WWSSTGTAWVSWASSA--LETSTQPATGATWKWHYAGSSRSPTLEA-TLTVSPFLAS	81	
Db	9	WRTRCSTCGRTWPSSARRURRST---ARWGMTL-----AASPACNAPTVTVAAPFGS	57	
QY	82	-LFVARVCRLRL-----CPYPKDSSTEPSNR-----VAMPSCPASLP	118	
Db	58	TLRATTLCRAITALDAANGSTLRCGIEPCPPTFESRVTLRSWEALTGPARKPRCPASSP	117	
QY	119	AQLMSSFRWFTCLP-----VTKLTLRPWMAACGARVKER	153	
Db	118	GML-----CMPNSAQCKSSSPSIAGAPPWPSSAGWKISRF	156	

Search completed: September 13, 2004, 14:23:55
Job time : 43.6316 secs

Db 4690 RRASSSTSHHTCSAGAPRSPSRSCSTTCRVPMSCGRSAACRTTASGPSRAICPRC 4749
Qy 95 --PYPKDSSTEPSNRV-----AWPSCPA-----SLPAQMSGPRWPTCLPVTKLTLR 140
Db 4750 STPAPTTWSATSTWTTTCGRSSGWPSPAPARWMPSPDPSTABASRCSCWPAAPLARSTSR 4809
Qy 141 PWWAACGAR 149
Db 4810 RWPASAPER 4818

RESULT 2

US-10-437-963-132079
; Sequence 132079, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132079
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_34082C.1.pap
US-10-437-963-132079

Query Match 13.7%; Score 125; DB 16; Length 437;
Best Local Similarity 25.0%; Pred. No. 0.013;
Matches 51; Conservative 7; Mismatches 64; Indels 82; Gaps 10;
Qy 11 PCWMPVCPMLPWRTW-----W-----WSSSSTAWVS 35
Db 11 PCCTSPAMPGRSWSCTPPTTGTCTPAQAGSSRSCWCTPPTTGTCTSPAWAGSSTWCS 70
Qy 36 W-----ASSALETSTOPATG-----ATWKMLHYAGSSR---ISPTLEATLTVSPF 78
Db 71 TPARGSSRNWCSTPTTGTCTSPAWAGSSTTWCSSTLARAGSSRSCWCTPT--TTGCTSPA 128
Qy 79 LASLRVARVCLRLCPYPKDSSTEPSW---RVAW-----PSCPASLP 118
Db 129 WAESSTT-----WCSTPPTTGTCTSPAWAGSSTTWCSSTLARAGSSRNWCSTTPTTGTCTSP 182
Qy 119 AQLMSGPRWPTCLPVTKLTLRPW 142
Db 183 AWAGSSTTWST--PARAGSSRW 204

RESULT 3

US-10-084-846A-4
; Sequence 4, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815

; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
; LENGTH: 19725
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 2: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: gat, Start position: nucleotide 2.
US-10-084-846A-4

Query Match 12.9%; Score 117.5; DB 15; Length 19725;
Best Local Similarity 29.8%; Pred. No. 2.4;
Matches 45; Conservative 13; Mismatches 48; Indels 45; Gaps 8;
Qy 24 WWSST-----AW-----VSWASSALETSTOPA---TGATW---TKWLHYAG 61
Db 12013 WRNPSSATVRSARAWCWRSTWRPCGGRWCSSARTTATPACRWGTG-RWPATRWGRPG 12071
Qy 62 SSRSPTLEATLTVSPFLASLRVARVCLRLCPYPKDSSTEPSRWRVAMPSCPASLPAL 121
Db 12072 SARRARWTEWTTPTSP-----TRPPRPSNAAAGAAVPPSWTAAPTGTATT 12117
Qy 122 MSSPRWPTCLPVTKLTLRPWWAACGARVKR 152
Db 12118 PERSRW-----GTTGTTPTP-----SASGVRT 12140

RESULT 4

US-10-437-963-132077
; Sequence 132077, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132077
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(278)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_34080C.1.pap
US-10-437-963-132077

Query Match 12.6%; Score 114.5; DB 16; Length 278;
Best Local Similarity 26.8%; Pred. No. 0.063;
Matches 51; Conservative 4; Mismatches 68; Indels 67; Gaps 10;
Qy 16 PCWLPWRTW-----WSSSSTAWV-----SW-----ASS 39
Db 53 PAWAGSSTWCLTTPPTTCTCTSPAWAGSSTTWCSSTPARAGSSRSCWCTTGTCTTTPAQAGSS 112
Qy 40 ALETSTQATGATWTKLHYAGSS---RISPTLEATLTVSPFLA-----SLRVAVC 88
Db 113 RWCSTPPTTGTCTSPAW---AGSSTTWCSTPPTTGTCT---SPVXAGSSSTTWSTWSTPARARS 166

```

QY      89 LRLCPYPKDSSTPSK---RVWPSCPASLPQALMSSPRWPTCLPVTKLTLPW--- 142
Db      167 SRWCSTPTTGCTSPAWAGSSTWCSTPAQ-----AGSSRWCSSTPTTGCTSPVWAGS 221
QY      143 ---WAACGAR 149
Db      222 STTWCSTEAR 231

```

RESULT 5

```

US-10-029-386-33554
; Sequence 33554, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33554
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008752.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: P11137, EVALUATION 6.00e-01
US-10-029-386-33554

```

```

Query Match      12.1%; Score 110.5; DB 14; Length 142;
Best Local Similarity 27.3%; Pred. No. 0.072;
Matches 36; Conservative 12; Mismatches 35; Indels 49; Gaps 8;

QY      12 CMVPCWL-PWRTWSSSTA-----WVSWASSALETST-OPATGA---TWTKLHYAGS 62
Db      41 CTTAPCSITSTRATWNTTSARALCWCRGASRPVTTTSPFGAASPTWTSWTRAGS 100
QY      63 SRISPTLEATLVSPFLASLRVARVCLLCPYPKDSSTPSMRVAV-----PSC 113
Db      101 G-----LCTPPTRTAT--SWSAGWTRTPSRSCGPGT 130
QY      114 PASLPALQMSSP 125
Db      131 PAT-PSAALARP 141

```

RESULT 6

```

US-10-437-963-168392
; Sequence 168392, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

```

```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168392
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(191)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66910C.1.pap
US-10-437-963-168392

```

```

Query Match      12.0%; Score 109.5; DB 16; Length 191;
Best Local Similarity 27.7%; Pred. No. 0.12;
Matches 44; Conservative 15; Mismatches 57; Indels 43; Gaps 5;

QY      14 MVPCLWRTWSSSTAWVSSALETSTQATGATWTKLHYAGSSRI--SPTLEA 71
Db      17 LIRIWLWSSPAMPFPSPAPATASAMCSSTPSSTPAC-----RGSATVASSPSSAA 69
QY      72 TLTVSPFLASLRVAR-----VCLRLCPYPKDSST 102
Db      70 TTTTTP---STRWCRRPCSSANGTSRRWXGRAGRTTPTATSTASCRRLACSPWMSST 126
QY      103 EPSRWVAMPSCFASLPALQMSSPRWPTCLPVTKLTLP 141
Db      127 SSS--TSTSTASSPATTGALRWPPATSATSTGTTTP 163

```

RESULT 7

```

US-10-425-114-46623
; Sequence 46623, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46623
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700423983_FLI.pap
US-10-425-114-46623

```

```

Query Match      11.8%; Score 108; DB 12; Length 217;
Best Local Similarity 23.1%; Pred. No. 0.18;
Matches 43; Conservative 17; Mismatches 50; Indels 76; Gaps 8;

QY      9 WLPCKMVPCWLPWRTW-----WSSSSTAWVSWASSAL 41
Db      7 WPEC-----WASSWDRPGAASRGSRPAPTRAPSRPSARCMARSCCSGWSRSGG 58
QY      42 ETST---OPATGATWTKLHYAGSSRIPTLEATLVSPFLASLRVARVCLRLCPYP 97
Db      59 WTSSITRPSFG-----PRGPALPRTSPTPATRATRAAGSSSTASAPAS 106
QY      98 XOSSTEPSWR----VAMP--SCPASLPALQMSSPRWPTCLPVTKLTLPWAAACGARVK 151
Db      107 RTASTSPTSRSSGLSWPSPNTPAS-----FCLPCSFSTSRP-----DRAR 147
QY      152 RFLQL 157

```

Db 148 RRSARL 153

RESULT 8

US-10-084-846A-3
; Sequence 3, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUEHLEWEG, AGNES
; APPLICANT: TREZGER, AXEL
; APPLICANT: BECHTOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 3
; LENGTH: 19695
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-3

Query Match 11.7%; Score 106.5; DB 15; Length 19695;
Best Local Similarity 22.9%; Pred. No. 20;
Matches 46; Conservative 16; Mismatches 50; Indels 89; Gaps 10;
QY 11 PCMMVPCWLPW-----RTWSSSSSTAWVS----- 35
Db 11303 PCSRTAAVFWPASPAPTRRSRSPARTWRCGASTTWAAGRS 11362
QY 36 -----WASSALETSTQATGATWTKWLYAGSS---RISPTLEATLVSPFLASLRVAR 86
Db 11363 AVKNCWSSS---TRPPTSGSTTRAVCGAGSTWPPDSSPCRCSSWTSPPASTRAT- 11418
QY 87 VCLRLCPYPKDSSTEPSWRVAV-----PSCPASLPAQLMSSPR-WW--PTCLPVTKLTL 139
Db 11419 -----ATRSGRWSVRWSSRAPPC-----CSPRSTWRRPTGSPTRSRST 11456
QY 140 R-----PWAAACGA 148
Db 11457 RAAGSRAPPRPSPWAAATGS 11477

RESULT 9

US-10-029-386-33373
; Sequence 33373, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33373
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL135838.2

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.69
; OTHER INFORMATION: SWISSPROT HIT: P13889, EVALUAE 6.30e+00
US-10-029-386-33373

Query Match 10.6%; Score 97; DB 14; Length 249;
Best Local Similarity 23.9%; Pred. No. 1.8;
Matches 51; Conservative 11; Mismatches 61; Indels 90; Gaps 12;
QY 17 CWLPWRTW-----WSSSSSTAWVS-----WASSALETS-----TOPATGATWT 54
Db 1 CWACTRAWARRSWGSPSTRWCSCGCRATPSGFWATTTRPSTSSWQVRRRAPSSASSAPWS 60
QY 55 KWLHYAGSSRI-----SPTLEATLVSPFLASLRVARVCLRLCP-----PYKDSSTE--- 103
Db 61 -WPRRGSCRTAORAPT-RARWTASRSTGTTRVCVASTGAWCPRCCVRRPASASTSSPM 118
QY 104 -----PS--WR-VAMPSCPASLP----- 118
Db 119 TLSRGRWAASRATACWCSCCWRVRAQASCFCSLPILWTWSSRGCRRTDCGAPRATAASW 178
QY 119 --AQLMSSPR-----WPTCLPVTKLTLRP 141
Db 179 TACTRATAPRAGASSHGGRPRCCCAPSPSTLPP 211

RESULT 10

US-10-425-114-38596
; Sequence 38596, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38596
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-050-D10_FLI1.pep
US-10-425-114-38596

Query Match 10.4%; Score 95; DB 12; Length 223;
Best Local Similarity 25.0%; Pred. No. 2.4;
Matches 32; Conservative 10; Mismatches 40; Indels 46; Gaps 6;
QY 12 CMWVPCWLPRTWSSSSSTAWVSASSALETSPTATGATWTKWLYAGSSRISPTLEA 71
Db 119 CRSCPC-RWQSPFCWTAPRRSSAWRR---QSRRP---SAMPASSRRGSRSGP--- 167
QY 72 TLTVSPFLASLRVARVCLRLCPYPKDSSTEPSWRVAVWPCPA-----SLPAQL 121
Db 168 -----CSRRLCFSPF-----ATCWRSCTSPSPAPSPPTSPSPAPS 202
QY 122 MSSPRWMP 129
Db 203 PSSPPWYP 210

RESULT 11

US-10-437-963-176482
; Sequence 176482, Application US/10437963

Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 176482
LENGTH: 265
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_74225C.1.pep
US-10-437-963-176482

Query Match 10.4%; Score 95; DB 16; Length 265;
Best Local Similarity 22.2%; Pred. No. 2.8;
Matches 40; Conservative 14; Mismatches 46; Indels 80; Gaps 7;

QY 53 WTKLHYAGSSRSPTLEAT-----LTVSPFLASLRVARYCL- 89
Db 54 WTSFARSSAASRPTPTASAPRAARGARPGNTARRRCRRASASPTAASASSRTRG 113
QY 90 RLICPPVPKDS-----TEPSRVVWPCSPALQOLMSSPR 126
Db 114 RSCPPPPAGSAPAGSCGTATTARTCSPTSPRPWWERS-PACPVSMSTTPSPS 172
QY 127 WMPTE-----LPVTKLTLRP-----AAAGARVKKR 153
Db 173 TSETCGAGRGSRATPASRRPRRCSSSPCPRPPTWRRWAKGGTASSRCAGRRR 232

RESULT 12
US-10-084-846A-6
Sequence 6, Application US/10084846A
Publication No. US2004006026A1
GENERAL INFORMATION:
APPLICANT: WEITNAUER, GABRIELE
APPLICANT: MUHLENWEG, AGNES
APPLICANT: TREFFZER, AXEL
APPLICANT: BECHTHOLD, ANDREAS
TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
FILE REFERENCE: 1974-005
CURRENT APPLICATION NUMBER: US/10/084,846A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: PCT/EP01/09815
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: DE 101 09 166.4
PRIOR FILING DATE: 2001-02-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: Patent In Ver. 3.2
SEQ ID NO 6
LENGTH: 19662
TYPE: PRT
ORGANISM: Streptomyces viridochromogenes
FEATURE:
OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-6

Query Match 10.3%; Score 94; DB 15; Length 19662;
Best Local Similarity 25.0%; Pred. No. 2.4e+02;
Matches 39; Conservative 13; Mismatches 60; Indels 44; Gaps 9;

QY 16 PCWLPWRTWWSSSTAWVSWASSALETSTQATGATWTKLHYAGSSRSPTLEATLV 75
Db 805 PC-RGWRRRWRTSAPW-----RPGTG--WPGWRGGRARRRPRAGSPSTY 849
QY 76 -----SPFLASLRVARYCLRLCPPY-----PKDSSTEPSRVVWPCSPALQ 119
Db 850 RWCWSPSRLSSR-----PRWNRGRGRGRSRWSSRP-WRPSWSSRRPWRPV 898
QY 120 QLMSSPRWPTCLPVTKLTLRPWAA--CGARVKKR 153
Db 899 RVTSMRWSTRCRPPRRRRRRRPPWECVRCRGPCHRR 934

RESULT 13
US-09-984-271-240
Sequence 240, Application US/09984271
Publication No. US20030040088A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: PZ030P1
CURRENT APPLICATION NUMBER: US/09/984,271
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 09/482,273
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: PCT/US99/15849
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,922
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,956
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 240
LENGTH: 145
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-271-240

Query Match 10.2%; Score 93; DB 10; Length 145;
Best Local Similarity 26.5%; Pred. No. 2.3;
Matches 36; Conservative 11; Mismatches 61; Indels 28; Gaps 5;

QY 12 CMVPCWLPWRTWWSSSTAWVSW-----ASSALETSTQATGATWTKLHYAGSSR 64
Db 12 CLWCPLWPAWPL--RCRPSANKWPPPLVGPAPKLGCSMTTQPTAVSWPCWL--MSSSL 67
QY 65 ISPTLEATLVSPFLASLRVARYCLLCPYPKDSSTEPSRVVWPCSPALQALMSS 124
Db 68 STACLAWTLTGLAREATRRAR-----SLSPWNCSARQVPPSPPHSGIGR 113
QY 125 PRW---WPTCLPVTKL 137
Db 114 RGWAHCHLTCLLVTLQ 129

RESULT 14
US-09-984-276-240
Sequence 240, Application US/09984276
Publication No. US20030017500A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: PZ030P1
CURRENT APPLICATION NUMBER: US/09/984,276
CURRENT FILING DATE: 2001-10-29
PRIOR APPLICATION NUMBER: 09/482,273
PRIOR FILING DATE: 2000-01-13
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/092,922

```

; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 240
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-276-240

```

RESULT 15
US-10-029-386-32972
; Sequence 32972, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32972
; SEQ ID NO 34288
; SEQ ID NO 34289

```

? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AL096708.11
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
? OTHER INFORMATION: SWISSPROT HIT: P27571, EVALU8 8.60e+00
? US-10-029-386-32972

```

Query Match	10.2%	Score 93;	DB 14;	Length 240;
Best Local Similarity	24.4%	Pred. No. 3.7;		
Matches 39;	Conservative 9;	Mismatches 54;	Indels 58;	Gaps 11;
Qy	8	AWLPCMMVPC-----WLP-----WPTWSSSSTAW-VSWASSALETSTQTATG	50	
Db	80	AWMTGFCVPCSRWRRAWMPGFCPCSWWRRAWMPGFCVCPWRRAWMPGFCVCP	139	
Qy	51	ATWTKWLHYAGSSRISPTLEATLTVSPPLASLRVA---RVCLRLLCPPYPKCSSTBPS	106	
Db	140	RWRRAW-----PGF-----VCPCSRWRRAWMPGFCV-----PWP-----PRW	173	
Qy	107	RYAW-PS---CPASLPAQLMSSRPWWPTCLPVTKLTLPWW	143	
Db	174	RRWWMPGFCVPC-----LPRWRRAWMPGFCVCPWW	205	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:10:50 ; Search time 45.4737 Seconds
(without alignments)
342.682 Million cell updates/sec

Title: US-10-019-219A-1

Perfect score: 912

Sequence: 1 TVVRLFLAWLPCMVPCWLP.....WAACGARVKRRFLQTLSTLSR 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	10.5	157	2 S73704	hypothetical prote
2	87.5	9.6	790	2 T34293	hypothetical prote
3	85.5	9.4	261	2 JQ0137	hypothetical 30.1k
4	84.5	9.3	210	2 H72460	hypothetical prote
5	83.5	9.2	384	2 G82670	general secretory
6	83.5	9.2	496	2 A54770	N-acetylglucosamin
7	83	9.1	332	2 C72310	conserved hypothet
8	83	9.1	377	2 F95948	probable acyltrans
9	81.5	8.9	254	2 B83286	hypothetical prote
10	81	8.9	984	2 T00326	hypothetical prote
11	81	8.9	1098	2 G70697	hypothetical prote
12	81	8.9	1522	2 T00028	probable arabinosy
13	80	8.8	610	2 S35049	brain-specific ang
14	79.5	8.7	660	2 JWM067	mucin JER57 - huma
15	79	8.7	38	2 S26903	chitinase (EC 3.2.
16	79	8.7	98	2 S12414	Ig heavy chain V r
17	79	8.7	1007	2 T01437	Ig heavy chain V r
18	78.5	8.6	784	2 JQ0317	hypothetical prote
19	78	8.6	3570	2 T45025	hypothetical 82K p
20	77.5	8.5	253	2 AC0810	mucin MUC5B, trach
21	77.5	8.5	283	1 S43189	probable sulfate t
22	77	8.4	118	2 A72474	hypothetical prote
23	77	8.4	145	2 S78055	hypothetical prote
24	77	8.4	146	2 S09711	Ig heavy chain pre
25	77	8.4	211	2 S73791	Ig heavy chain V r
26	77	8.4	1513	2 A54895	hypothetical prote
27	77	8.4	1859	1 A34092	mucin 2, intestina
28	77	8.4	1862	2 T29959	DNA-directed RNA p
29	76.5	8.4	99	2 S12412	DNA-directed RNA p
					Ig heavy chain V r

30 76.5 8.4 118 2 A26340 Ig heavy chain pre
31 76.5 8.4 130 2 S31690 Ig heavy chain V r
32 76.5 8.4 290 1 BVEBC2 cys2 protein - Sal
33 76.5 8.4 304 2 S39892 rnfD protein - Rho
34 76.5 8.4 466 2 AD3477 malate dehydrogena
35 76.5 8.4 516 2 S19586 N-methyl-D-asparta
36 76.5 8.4 580 2 T43481 probable mucin DKF
37 76.5 8.4 825 2 S75173 hypohetical prote
38 76.5 8.4 947 2 A12787 valyl-tRNA synthet
39 76.5 8.4 947 2 B97567 valyl-tRNA synthet
40 76.5 8.4 1711 1 A55148 protein-tyrosine-p
41 76 8.3 98 2 S26905 Ig heavy chain V r
42 76 8.3 157 2 S9805 hypothetical prote
43 76 8.3 198 2 T35915 hypothetical prote
44 76 8.3 235 2 I38440 flt3 ligand - huma
45 76 8.3 296 2 JC7268 paired-type homeod

ALIGNMENTS

RESULT 1

S73704

hypothetical protein H08_orf157a - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342

C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000

C:Accession: S73704

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73704

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-157 <HIM>

A:Cross-references: EMBL:AE000036; GB:U00089; NID:91674053; PIDN:AAB96026.1; FID:91674051

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf157a

Query Match

Best Local Similarity 10.5%; Score 96; DB 2; Length 157;

Matches 36; Conservative 19; Mismatches 62; Indels 34; Gaps 8;

Qy 9 WLPCCMVPCWLP-----W-RTWWSSTSTAWVSW-----ASSALETSTQ----- 46

Db 4 WMSC-APPYTPHTNSWTESGMDRTSWRWSAQRWSGSKIVRANKALRYMAKTKMLPV 62

Qy 47 ---PATGATWTKW-----LHYAGSSRISPTLEATLTVSPFLASLRVARVCLRLCPVPK 98

Db 63 LIPSPNKPKYSKLAINQELHLTPHKKTSPTSSSLKRPFGPRGVLNARLSWR--CPTLSR 120

Qy 99 DSSTESWRVAVFSCPASLPAQLMSSPRWMP 129

Db 121 KVRV-PTIKVPMVRAPSTKESKTSNNPWP 150

RESULT 2

T34293

hypothetical protein F49E10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000

C:Accession: T34293

R:Miller, N.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid F49E10.

A:Reference number: Z21500

A:Accession: T34293

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-790 <MIL>

Query Match 9.3%; Score 84.5; DB 2; Length 210;
Best Local Similarity 25.0%; Pred. No. 1.4;

C:Accession: A54770: S37024

Query Match 9.3%; Score 84.5; DB 2; Length 210;
Best Local Similarity 25.0%; Pred. NO.1.4;

Genomics 21, 18-26, 1994

A:Title: Molecular cloning, cDNA analysis, and localization of a monomer of the N-acetyl
A:Reference number: A54770; MUID:94375011; PMID:8088785

A:Accession: A54770

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-496 <BLA>

A:Cross-references: EMBL:X72018

C:Genetics:

A:Gene: GDB:NAGRI

A:Cross-references: GDB:250465; OMIM:160994

A:Map position: 19p13.2-19p13.2

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
C:Keywords: calcium binding; lectin; thyroid gland; transmembrane protein

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-496/Product: N-acetylglucosamine receptor 1 #status predicted <MAT>

F:43-111/Domain: ribonucleoprotein repeat homology <RRM2>

F:230-236/Region: glycine-rich

F:441-460/Domain: transmembrane #status predicted <TM>

F:488-495/Region: coated-pit mediated internalization signal

Query Match

Best Local Similarity 9.2%; Score 83.5; DB 2; Length 496;

Matches 27; Conservative 13; Mismatches 35; Indels 39; Gaps 6;

QY 14 MVPCLPWRWWWSSSTAWVSWASSALETSTQPATGATWTKLHYAGSSRISPTLEATL 73

Db 391 MPFPAWSAWPTIW--SGWASAWAPT-----SSAWASAWPTASSAWAPP----- 435

QY 74 TVSPFLASLRVARVCLRLCPYPKDSSTPSNRVAVPSCFASLPALQMLGSPRW 127

Db 436 WARDWALAL-----SAW--ANFWAVAVPALTIVPS-RW 465

RESULT 7

C72310

conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: C72310

R:Nelson, K.E.; Stewart, A.M.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: C72310

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332 <ARN>

A:Cross-references: GB:AE001760; GB:AE000512; NID:94981510; PIDN:AAD36061.1; PID:9498152
A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TW0982

Query Match

Best Local Similarity 9.1%; Score 83; DB 2; Length 332;

Matches 30; Conservative 20; Mismatches 42; Indels 20; Gaps 6;

QY 30 STAWVSWASSALET-STQPATGATWTKWLH-----YAGSSRISPTLEATLTVS 76

Db 104 TTAWFAALAYALTAAHATKGLFSGWIKWLSNYTVTSNTNPVA--PKTGTTATVLTGIS 161

QY 77 PFLASLRVARVCLRLCPYPKDSSTPSNRVAVPSCFASLPALQMLGSPRW 128

Db 162 PWIAS--IIFLALMLWYAFGVKNKORPS-KLNN--IAASVLIALIAPIAW 208

RESULT 8

F95948

probable acyltransferase, possibly surface-saccharide specific acetyltransferase protein
C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: F95948

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11461431

A:Accession: F95948

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-377 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49254.1; PID:g15140740; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.

Science 293, 568-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

C:Contents: annotation

C:Genetics:

A:Gene: SMB21188

A:Genome: plasmid

Query Match

Best Local Similarity 9.1%; Score 83; DB 2; Length 377;

Matches 36; Conservative 11; Mismatches 57; Indels 54; Gaps 7;

QY 18 WLPW---RTWWSSSTAWVSWASSALETSTQPATGAT-----WTKWLHYAG 61

Db 120 WTPVWLCHTWSLSIEEFYLVWPLVLLVPRRSVAGVGVVCSLAYRFYW----- 171

QY 62 SSRISPTLEATLTVSPFLASLRVARVCLRLCPYPKDS-----STEPSNRVAVPSCP 114

Db 172 -----PLTGTPLA-----RDLLPASM DALAVGALLAARPSRWSGNFAWA 212

QY 115 --ASLPAQLMSSPRWPTCLPVTKLTLPWPWAAAGARV 150

Db 213 KLSWMLPLSLASLCILVMSK--PVAMTPVVAWFAWMTGLEV 248

RESULT 9

B83286

hypothetical protein PA2884 [imported] - Pseudomonas aeruginosa (strain PAO1)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: B83286

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: B83286

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-254 <STO>

A:Cross-references: GB:AE004714; GB:AE004091; NID:g9948965; PIDN:AAG06272.1; GSPDB:GN001
A:Experimental source: strain PAO1

C:Genetics:

A:Gene: PA2884

Query Match

Best Local Similarity 8.9%; Score 81.5; DB 2; Length 254;

Matches 37; Conservative 14; Mismatches 62; Indels 47; Gaps 7;

QY 2 VTRLFLAWLPCMVPWCLPWRTWWSSTAWVSWASSAL-----PVSLAASALIAHYLPGLPLRFPWDLA 80

Db 30 IVLLYIIWV--TYAERWLPTLLWV-----PVSLAASALIAHYLPGLPLRFPWDLA 80

QY 44 STQPATGATWTKWLHYAGSSRISPTLEATLTVSPF--LASLRVARV--CLRLCPYPKPD 99

Db 81 SASPPTQMTWLPW----DKALVALTLTLLAWLRRPKQPLVSLDITALAFCLTFVFPVPLLSI 136

628 FLAALFFLLALCWATTNGWYVSSGVGFSPNSAMPKIDGTVSTIFFALFAIAAGYAALWH 687

QY 59 YAGSRSISPTLEALTIVSP-----FLASLRVARVCLRLCCPPYPKDS 101
Db 688 FAPRGAGEGLIRALTITAPVPIVAGFMAAVFVASMVAGIV-RQYPTYSN 735

RESULT 12
T00028
brain-specific angiogenesis inhibitor 3 - human
N;Alternate names: BAI3 protein
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00028
R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogetnet. Cell Genet. 79, 103-108, 1997
A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-
A;Reference number: Z14066; MUID:98194217; PMID:9533023
A;Accession: T00028
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-1522 <SHI>
A;Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BAA25363.1; PID:g3021701
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI3
A;Cross-references: GDB:9838090; OMIM:602684
A;Map position: 6q12-6q12
C;Superfamily: thrombospondin type 1 repeat homology
F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 8.9%; Score 81; DB 2; Length 1522;
Best Local Similarity 21.6%; Pred. No. 22;
Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;

QY 11 PCMMVPC-----WLPRTWWS-----SSTAWSVASSALETST 45
Db 391 PCNIALCPVDGQWQWSSQCSVTCSNGTQORSRQCTAAAHGSGECRGPWAESRECVNP 450
QY 46 QPATGATWTKWLHYAGSS-----RISPTLEALTIVSPFLASLRVARVCLRLCCP-P 95
Db 451 ECTANGQNWQHWGWSCKSCSDGGWERRINTCOGAVITGQCCGTGEVRCSEORCAP 510

QY 96 Y---PKDSSTEPSWR-----VAMPSCP 114
Db 511 YEICPEDYLMVMWVKRTPTAGDLAFNQCP 538

RESULT 13
S35049
mucin JER57 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 10-Dec-1993 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: S35049; S37594
R;Dufoss, J.; Porchet, N.; Audie, J.P.; Guyonnet Duperat, V.; Laine, A.; van-Seuningen,
Biochem. J. 293, 329-337, 1993
A;Title: Degenerate 87-base-pair tandem repeats create hydrophilic/hydrophobic alternati
A;Reference number: S35047; MUID:93343858; PMID:7916618
A;Accession: S35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-610 <DUF>
R;Aubert, J.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37593
A;Accession: S37594
A;Molecule type: mRNA
A;Residues: 1-20, 'W', 22-610 <AUB>
A;Cross-references: EMBL:X74955

Query Match 8.8%; Score 80; DB 2; Length 610;
Best Local Similarity 23.9%; Pred. No. 11;
Matches 16; Conservative 16; Mismatches 19; Indels 16; Gaps 2;

QY 10 LPCMVPCLWPRTWSSSTAWVSWASSALETSTQATGATWTKWLHYAGSSRISPTL 69
Db 3 LECRASPAWCPGWSARSTNAATLAWGA-----GTVSRW-----GSSRCASTM 46
QY 70 EATLTVS 76
Db 47 KSVCSAA 53

RESULT 14

JW0067
chitinase (EC 3.2.1.14) A - Emericella nidulans
N:Alternate names: chia
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 11-Jan-2002
C:Accession: JW0067
R:Takaya, N.; Yamazaki, D.; Horiuchi, H.; Ohta, A.; Takagi, M.
BioSci. Biotechnol. Biochem. 62: 60-65, 1998
A:Title: Cloning and characterization of a chitinase-encoding gene (chia) from Aspergillus
A:Reference number: JW0067; MUID:98162139; PMID:9501518
A:Accession: JW0067
A:Molecule type: mRNA
A:Residues: 1-660 <TAK>
A:Cross-references: DDBJ:D87895; NID:G2821948; PID:G2828335
C:Comment: This enzyme hydrolyzes chitin at Delta-1,4 bonds between N-acetylglucosamine
C:Genetics: chia
A:Gene: chia
C:Keywords: glycosidase; hydrolase

Query Match 8.7%; Score 79.5; DB 2; Length 660;
Best Local Similarity 25.3%; Pred. No. 13;
Matches 37; Conservative 15; Mismatches 39; Indels 55; Gaps 6;

QY 27 SSSSTAWVSWASSALETSTQ-----PATGATWTKWLHYAGSSRISPTLEAT 72
Db 490 STSSTRAVSETSTHISTSTSSGPTSLTGSSTSVPATSS-----VPSSAIFSPSTPV 542

QY 73 LTVSP-----FLASLRVARVCLRLCPYPKDSST-----EPSRWVAMPSCP 114
Db 543 ISETPRPVTSSTSSSTFVSSTSTDC-----SESSTAIGHSSSIWETPSASTP 593

QY 115 ASLPA-----QLMSSPRWPTCLP 133
Db 594 AASFSTSPETTKTLTVFRRAPPCLP 619

RESULT 15

S26903
Ig heavy chain V region (DP-68 / 4.13) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S26903; S12413
R:Tomlinson, I.M.; Walter, G.; Marks, J.D.; Llewellyn, M.B.; Winter, G.
J. Mol. Biol. 227: 776-798, 1992
A:Title: The repertoire of human germline V(H) sequences reveals about fifty groups of V
A:Reference number: S26885; MUID:93021117; PMID:1404388
A:Accession: S26903
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <TOM>
A:Cross-references: EMBL:Z12368; NID:G32954; PIDN:CAA78238.1; PID:G32955
R:Sanz, I.; Kelly, P.; Williams, C.; Scholl, S.; Tucker, P.; Capra, J.D.
EMBO J. 8: 3741-3748, 1989
A:Title: The smaller human V(H) gene families display remarkably little polymorphism.
A:Reference number: S09421; MUID:90059975; PMID:2511001
A:Accession: S12413
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-98 <SAN>
A:Cross-references: EMBL:X56357
A:Note: designated 4.13

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 8.7%; Score 79; DB 2; Length 98;
Best Local Similarity 29.4%; Pred. No. 2;
Matches 15; Conservative 13; Mismatches 17; Indels 6; Gaps 1;

QY 26 WSSSTAWVSWASSALETSTQATGATWTKWLHYAGSSRISPTLEATLTVS 76
Db 27 YSISSSNWGWIRQ-----PPGKLEWIGVYIYSGSTYNNPSLSKSRVTMS 71

Search completed: September 13, 2004, 14:23:14
Job time : 57.4737 secs

Result No.	Query No.	Score	Query			ID	Description
			Match	Length	DB		
1	1	96	10.5	157	1	YE63_MYCPN	P75320 mycoplasma
2	2	88	9.6	1224	1	AT16_HUMAN	Q8te57 homo sapien
3	3	85.5	9.4	890	1	WFS1_MOUSE	P56695 mus musculu
4	4	81.5	8.9	653	1	SF01_MOUSE	Q64213 mus musculu
5	5	81	8.9	1098	1	EMBB_MYCTU	P72030 mycobacteri
6	6	81	8.9	1522	1	EA13_HUMAN	Q60242 homo sapien
7	7	80.5	8.8	1595	1	LTBL_HUMAN	Q14766 homo sapien
8	8	78.5	8.6	784	1	YAV2_XANCV	P14728 xanthomonas
9	9	78.5	8.6	1065	1	EMBB_MYCAV	P71486 mycobacteri
10	10	78.5	8.6	1139	1	M2A2_HUMAN	P36641 homo sapien
11	11	78.5	8.6	1527	1	ARHB_RAT	Q9es67 rattus norv
12	12	78	8.6	5703	1	M05B_HUMAN	Q9hc84 homo sapien
13	13	77.5	8.5	253	1	CYSZ_SALTI	Q8z4w3 salmonella
14	14	77	8.4	211	1	YD71_MYCPN	P75410 mycoplasma
15	15	77	8.4	487	1	HEP_DROME	Q23977 drosophila
16	16	77	8.4	1513	1	MUC2_RAT	Q62635 rattus norv
17	17	77	8.4	1852	1	RBPI_CAELT	P16356 caenorhabdi
18	18	76.5	8.4	253	1	CYSZ_SALTY	P12673 salmonella
19	19	76.5	8.4	890	1	WFS1_HUMAN	Q76024 homo sapien
20	20	76.5	8.4	1711	1	PTPV_RAT	Q64612 rattus norv
21	21	76	8.3	157	1	UL42_HCWA	P16815 human cytom
22	22	76	8.3	235	1	FL3L_HUMAN	P49771 homo sapien
23	23	76	8.3	635	1	TPOR_HUMAN	P40238 homo sapien
24	24	75.5	8.3	886	1	PLSB_XANCP	Q8p3e3 xanthomonas
25	25	75.5	8.3	2845	1	APC_MOUSE	Q61315 mus musculu
26	26	75	8.2	725	1	AGA1_YEAST	P32323 saccharomyc
27	27	74.5	8.2	218	1	CSF2_RAT	Q9j1h5 rattus norv
28	28	74.5	8.2	546	1	AAAS_HUMAN	Q9nr99 homo sapien
29	29	74.5	8.2	951	1	HEX_ADE05	P04133 human adeno
30	30	74.5	8.2	967	1	HGX_ADE02	P03277 human adeno
31	31	74	8.1	348	1	HFE_HUMAN	Q30201 homo sapien
32	32	74	8.1	348	1	HFE_PANTR	P60018 pan troglod
33	33	74	8.1	1776	1	POLR_OYMW	P20127 ononis yell

AC Q8TE57; Q8IVE2; 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-16 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 16) (ADAM-TS 16) (ADAM-TS16).
 GN ADAMTS16 OR KIAA2029.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 RA Lopez-Otin C.,
 RT "Cloning, expression analysis, and structural characterization of
 RT seven novel human ADAMTSs, a family of metalloproteinases with
 RT disintegrin and thrombospondin-1 domains.";
 RL Gene 283:49-62(2002).
 RN [2]
 RP SEQUENCE OF 204-1224 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human.";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -! COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -! SURCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -! ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8TE57-1; Sequences=Displayed;
 CC Name=2;
 CC IsoId=Q8TE57-2; Sequences=VSP_007664, VSP_007665;
 CC TISSUE SPECIFICITY: Expressed in fetal lung and kidney and in
 CC adult prostate and ovary.
 CC -! DOMAIN: The spacer domain and the TSP type-1 domains are important
 CC for a tight interaction with the extracellular matrix (By
 CC similarity).
 CC -! PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -! SIMILARITY: Belongs to peptidase family M12B.
 CC -! SIMILARITY: Contains 1 disintegrin-like domain.
 CC -! SIMILARITY: Contains 6 TSP type-1 domains.
 CC -! SIMILARITY: Contains 1 PLAC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ315734; CAC86015.1; -
 CC EMBL; AB05949; BAC23125.1; -
 CC MEROPS; M12.026; -
 CC Genew; HGNC:17108; ADAMTS16.
 CC MIM; 607510; -
 CC InterPro; IPR001762; Disintegrin
 CC InterPro; IPR001818; Pept_M12B_M12B.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR001590; Peptidase_M12B.
 CC InterPro; IPR002870; Peptidase_M12B_N.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR008085; TSP_1.
 CC Pfam; PF01562; Pept_M12B_propep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC Pfam; PF00090; tsp_1; 5.
 CC PRINTS; PR01705; TSP1REPEAT.
 CC SMART; SM00209; TSP1; 6.
 CC PROSITE; PS50215; ADAM_MEPRO; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
 DR PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
 DR PROSITE; PS0900; PLAC; 1.
 DR PROSITE; PS0092; TSP1; 5.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 Repeat; Extracellular matrix; Alternative splicing.
 FT SIGNAL 1 24
 FT PROPEP 25 279
 FT CHAIN 280 1224
 FT DOMAIN 280 495
 FT DOMAIN 496 585
 FT DOMAIN 586 641
 FT DOMAIN 642 746
 FT DOMAIN 747 873
 FT DOMAIN 874 922
 FT DOMAIN 927 987
 FT DOMAIN 988 1048
 FT DOMAIN 1051 1115
 FT DOMAIN 1127 1181
 FT DOMAIN 1186 1223
 FT SITE 249 249
 FT METAL 433 433
 FT ACT SITE 434 434
 FT METAL 437 437
 FT METAL 443 443
 FT CARBOHYD 156 156
 FT CARBOHYD 310 310
 FT CARBOHYD 741 741
 FT CARBOHYD 780 780
 FT CARBOHYD 835 835
 FT CARBOHYD 905 905
 FT CARBOHYD 935 935
 FT VARSPLIC 1063 1072
 FT VARSPLIC 1073 1224
 FT CONFLICT 284 284
 FT SEQUENCE 1224 AA; 136186 MW; 572E72129E1401F8 CRC64;
 SQ
 Query Match 9.6%; Score 88; DB 1; Length 1224;
 Best Local Similarity 25.5%; Pred. No. 2.1;
 Matches 37; Conservative 18; Mismatches 56; Indels 34; Gaps 7;
 QY 12 CMWPCWLPRTWVWSSSTAWVSWASSALETSTQATGATWTKWLYHAGSSRISPTLEA 71
 Db 1042 CILQCHKPKKQWLVSV---AW-SQSVTCERTQKFLKCAEKYV--SGKYRELASKKC 1095
 QY 72 TLTVSPFLASLRVARVCLRLCPYPKDSSTEPSWRVAMPSCPASLPALMSSFRWPTC 131
 Db 1096 SHLPKP---SLELERACAPLPCPRHPFFAAAGPS-RGSWFAFP-----WSQC 1138
 QY 132 LPVTKLTLRPWWAACGARYKRFLO 156
 Db 1139 T-----ASCGGVQTRSVQ 1152
 RESULT 3
 ID WFS1_MOUSE STANDARD; PRT; 890 AA.
 AC P56695; Q92276;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Wolframin.
 GN WFS1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC      EMBL; X80159; CAA56440.1; -
CC      EMBL; X85802; CAA59797.1; -
CC      EMBL; Y08907; CAA70113.1; -
CC      EMBL; Y12838; CAA73359.1; -
CC      EMBL; BC009091; AAH09091.1; -
CC      EMBL; BC055370; AAH55370.1; -
CC      MGD; MGI:1095403; Zfp162.
CC      InterPro; IPR004087; KH dom.
CC      InterPro; IPR004088; KH type 1.
CC      InterPro; IPR001878; ZnF_CCHC.
CC      Pfam; PF00013; KH; 1.
CC      Pfam; PF00098; ZF_CCHC; 1.
CC      PRINTS; PR00939; C2HCZNFINGER.
CC      SMART; SM00322; KH; 1.
CC      SMART; SM00343; ZnF_C2HC; 1.
CC      PROSITE; PSS0084; KH_TYPE_1; 1.
CC      PROSITE; PSS0158; ZF_CCHC; 1.
CC      mRNA processing; mRNA splicing; Transcription regulation; Repressor;
KW      Nuclear protein; Spliceosome; RNA-binding; Metal-binding; Zinc;
KW      Zinc-finger; Phosphorylation; Alternative splicing.
FT      DOMAIN 15 19 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      DOMAIN 141 222 KH.
FT      ZN_FING 277 296 CCHC-TYPE.
FT      DOMAIN 324 585 PRO-RICH.
FT      MOD_RES 20 20 PHOSPHORYLATION (by PKG) (BY SIMILARITY).
FT      VARSPLIC 529 653 TTTTTSAGTGPWPQQQAAAASFGTFQMGNTMVP
FT      PRCVQPLPGAPPTPTCTICLLSPNSLSPNR
FT      RIPPRGSDGSHESDFRPLVTLPGRQQRQWWTGFGK
FT      AA -> SLPAAMARARVTRFAHW (in isoform
FT      CM17E).
FT      /FTID=VSP_050424.
FT      TCSIECLLSPNSLCLSPNRAARIPRPGSDGSHESD
FT      FRPLVTLPGRQQRQWWTGFGKAA -> PPPPPGSA
FT      MYAPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
FT      PQN (in isoform 3).
FT      /FTID=VSP_008840.
FT      E -> K (IN REF. 3).
FT      P -> S (IN REF. 3); AAH55370).
FT      P -> L (IN REF. 3); AAH09091).
FT      N -> R (IN REF. 1).
FT      SEQUENCE 653 AA; 70408 MW; 8228DE3E79AB1BFC CRC64;
FT      QUERY MATCH 8.9%; Score 81.5; DB 1; Length 653;
FT      Best Local Similarity 21.4%; Pred. No. 4.3;
FT      Matches 30; Conservative 19; Mismatches 58; Indels 33; Gaps 4;
QY      27 SSSTAWVSSALETSTQATGA-TTWKHLHYGSSRISPTLEATLTVSPFLASLRV 84
DB      516 SMASTPLPQQNTTTTTSAGTGPWPQQQAAAASFGTFQMGNTMVPVLPVQV- 574
QY      85 ARVCLRLCPYPKDSSTEPSRWVAPSCPASLPAQLMSSSPRW----- 127
DB      575 -----PPLPGAPPTPTCTICLLSPNSLCLSPNRAARIPRPGSDGSHES 625
QY      128 -WP---TCLPVTKLTPRW 143
DB      626 DFRPLVTLPGRQQRQWFW 645
RESULT 5
ID      EMBL MYCTU STANDARD; PRT; 1098 AA.
AC      P72030; P72061;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)

```

```

DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Probable arabinosyltransferase B (EC 2.4.2.-).
GN      EMBL OR RV3795 OR MT3902 OR MTCY13D12.29.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=97287037; PubMed=9142129;
RA      Telenti A., Philipp W.J., Sreevatsan S., Bernasconi C.,
RA      Stockbauer K.E., Wiele B., Musser J.M., Jacobs W.R. Jr.;
RT      "The emb operon, a gene cluster of Mycobacterium tuberculosis involved
RT      in resistance to ethambutol.";
RL      Nat. Med. 3:567-570(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekaiia P.,
RA      Gordon S.V., Eiglmeier K., Gas S., Chillingworth T., Connor R.,
RA      Badcock K., Basham D., Brown D., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA      Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RA      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence.";
RL      Nature 393:537-544(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CDC 1551 / Oshkosh;
RX      MEDLINE=22206494; PubMed=12218036;
RA      Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA      Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA      Kolonay J.F., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA      Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA      Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA      "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT      laboratory strains.";
RL      J. Bacteriol. 184:5479-5490(2002).
RN      [4]
RP      VARIANTS EMB RESISTANT LEU-306; ILE-306; VAL-306 AND VAL-330.
RX      MEDLINE=97400246; PubMed=9257740;
RA      Sreevatsan S., Stockbauer K.E., Pan X., Kreiswirth B.N.,
RA      Moghazeh S.B., Jacobs W.R. Jr., Telenti A., Musser J.M.;
RT      "Ethambutol resistance in Mycobacterium tuberculosis: critical role of
RT      embB mutations.";
RL      Antimicrob. Agents Chemother. 41:1677-1681(1997).
RN      [5]
RP      VARIANTS EMB RESISTANT.
RX      MEDLINE=20106977; PubMed=10639358;
RA      Ramasamy S.V., Amin A.G., Goekseel S., Stager C.B., Dou S.-J.,
RA      El Sahly H., Moghazeh S.B., Kreiswirth B.N., Musser J.M.;
RT      "Molecular genetic analysis of nucleotide polymorphisms associated
RT      with ethambutol resistance in human isolates of Mycobacterium
RT      tuberculosis.";
RL      Antimicrob. Agents Chemother. 44:326-336(2000).
RN      [6]
RP      VARIANTS EMB RESISTANT LEU-306; ILE-306 AND VAL-306.
RX      MEDLINE=21106283; PubMed=11162078;
RA      Rinder H., Mieskes K.T., Tortoli E., Richter E., Casal M., Vaguero M.,
RA      Cambau E., Feldmann K., Loeschner T.;
RT      "Detection of embB codon 306 mutations in ethambutol resistant
RT      Mycobacterium tuberculosis directly from sputum samples: a low-cost,
RT      rapid approach.";
RL      Mol. Cell. Probes 15:37-42(2001).
CC      -!- FUNCTION: Arabinosyl transferase responsible for the
CC      polymerization of arabinose into the arabinan of arabinogalactan.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC      -!- MISCELLANEOUS: This is one of the target of the anit-tuberculosis

```

drug ethambutol [(S,S')-2,2'-(ethylenediimino)di-1-butanol; EMB].
 EMB is a first-line drug used to treat tuberculosis. EMB inhibits
 the transfer of arabinogalactan into the cell wall.
 -!- SIMILARITY: BELONGS TO THE EMB FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; Z80343; CAB02474.1; -.
 EMBL; U68480; AAC45281.1; -.
 EMBL; AE007183; AAK48268.1; -.
 PIR; G70697; G70697.
 TIGR; MT3902; -.
 TubercuList; RV3795; -.
 InterPro; IPR007680; Arab transf.
 Pfam; PF04602; arab transf; 1.
 Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 Antibiotic resistance; Complete proteome.
 TRANSMEM 28 50
 TRANSMEM 217 239
 TRANSMEM 271 293
 TRANSMEM 402 419
 TRANSMEM 434 456
 TRANSMEM 472 494
 TRANSMEM 541 558
 TRANSMEM 570 587
 TRANSMEM 597 619
 TRANSMEM 626 648
 TRANSMEM 663 685
 TRANSMEM 698 720
 TRANSMEM 297 297
 TRANSMEM 306 306
 TRANSMEM 306 306
 TRANSMEM 306 306
 TRANSMEM 328 328
 TRANSMEM 328 328
 TRANSMEM 330 330
 TRANSMEM 334 334
 TRANSMEM 406 406
 TRANSMEM 406 406
 TRANSMEM 406 406
 TRANSMEM 497 497
 TRANSMEM 497 497
 TRANSMEM 745 745
 TRANSMEM 959 959
 TRANSMEM 1000 1000
 TRANSMEM 1024 1024
 TRANSMEM 773 774
 TRANSMEM 1098 AA; 118020 SW -> FL (IN REF. 1).
 SSEQUENCE 1098 AA; 118020 SW -> FL (IN REF. 1).
 Query Match 8.9%; Score 81; DB 1; Length 1098;
 Best Local Similarity 26.68; Pred. No. 8.1;
 Matches 29; Conservative 16; Mismatches 50; Indels 14; Gaps 3;
 QY 6 FLAWLPCMVPCMLPWRWTWSSSTAWSVASSALETSTQP-----ATGATWTKWLH 58
 Db 628 FLAALFFLLALCWATTNGWVSSYGVFPNSAMPKIDGIVTSTFFALFAIAGYAALWLH 687
 QY 59 YAGSSRSIPLEATLTSP-----FLASLRVARVCLRLCPYPKQSS 101
 Db 688 FAPRGAGEGLRLATTPAPVIVAGFVAAVFVSMVAGIV-RQYPTYSN 735

RESULT 6
 BAI3_HUMAN
 ID_BAI3_HUMAN STANDARD; PRT; 1522 AA.
 AC O60242; O60297;
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brain-specific angiogenesis inhibitor 3 precursor.
 GN BAI3 OR KIAA0550.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plrimates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98194217; PubMed=9533023;
 RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
 RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
 RL to brain-specific angiogenesis inhibitor 1 (BAI1).";
 RL Cytogenet. Cell Genet. 79:103-108(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RL The complete sequences of 100 new cDNA clones from brain which can
 RN code for large proteins in vitro.";
 RN DNA Res. 5:31-39(1998).
 CC [3]
 RP SEQUENCE FROM N.A., AND REVISIONS TO 643-665 AND C-TERMINUS.
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RL curation of 330 KIAA cDNA clones.";
 RL DNA Res. 9:99-106(2002).
 CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
 CC SUPPRESSION OF GLOBLASTOMA.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
 CC HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL
 CC LINES.
 CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
 CC -!- SIMILARITY: Contains 1 CUB domain.
 CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 GPS domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 EMBL; AB005299; BAA25363.1; -.
 EMBL; AB011122; BAA25476.2; ALT_INIT.
 PIR; T00028; T00028.
 Genew; HGNC:945; BAI3.
 MIM; 602684; -.
 InterPro; IPR000859; CUB.
 InterPro; IPR000832; GPCR secretin.
 InterPro; IPR001879; hormn_receptor.
 InterPro; IPR000203; PKD_cys_rich.
 InterPro; IPR000884; TSP1.
 Pfam; PF00002; 7tm_2; 1.
 Pfam; PF01825; GPS; 1.
 Pfam; PF02793; HRM; 1.
 Pfam; PF00030; tsp_1; 4.
 SMART; SMC0303; GPS; 1.
 SMART; SMC0008; HORMR; 1.
 SMART; SMC0209; TSP1; 4.
 PROSITE; PS01180; CUB; 1.
 PROSITE; PS50221; GPS; 1.
 PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
 PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.

DR PROSITE; PS50227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECP_F2_4; 1.
 DR PROSITE; PS50092; TSP1; 4.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 1522
 FT DOMAIN 25 880
 FT TRANSMEM 881 901
 FT DOMAIN 902 910
 FT TRANSMEM 911 931
 FT DOMAIN 932 939
 FT TRANSMEM 940 960
 FT DOMAIN 961 981
 FT TRANSMEM 982 1002
 FT DOMAIN 1003 1023
 FT TRANSMEM 1024 1044
 FT DOMAIN 1045 1098
 FT TRANSMEM 1099 1119
 FT DOMAIN 1120 1125
 FT TRANSMEM 1126 1146
 FT DOMAIN 1147 1522
 FT DOMAIN 30 159
 FT DOMAIN 291 343
 FT DOMAIN 345 358
 FT DOMAIN 400 453
 FT DOMAIN 455 508
 FT DOMAIN 816 868
 FT DOMAIN 942 945
 FT DOMAIN 1173 1176
 FT CARBOHYD 51 51
 FT CARBOHYD 54 54
 FT CARBOHYD 82 82
 FT CARBOHYD 105 105
 FT CARBOHYD 241 241
 FT CARBOHYD 337 337
 FT CARBOHYD 418 418
 FT CARBOHYD 540 540
 FT CARBOHYD 625 625
 FT CARBOHYD 779 779
 FT CARBOHYD 812 812
 FT CARBOHYD 828 828
 FT CARBOHYD 937 937
 SQ SEQUENCE 1522 AA; 171490 MW; D22D0A5D4BB62502 CRC64;
 Query Match 8.9%; Score 81; DB 1; Length 1522;
 Best Local Similarity 21.6%; Pred. No. 11;
 Matches 32; Conservative 13; Mismatches 59; Indels 44; Gaps 6;
 QY 11 PCMVPC-----WLPRTWWS-----SSTAWVSWSSALETST 45
 DB 391 PCNIALCPVDGQWQWSSWSSQCSVTCNGTQQRQCTAAAHGGSGCRGFWAESRCYNP 450
 QY 46 OPATGATWTKWLIHYAGSS-----RISPTLEATLTVSPFLASLRVARVCLRLICP-P 95
 DB 451 ECTANGCQWQWGHWSGCSKSCDGGWRRRTCCGAVITGQCQCGTGEVRCSEQRCPAP 510
 QY 96 Y----PKDSSTEPSWR-----VAMPSCP 114
 DB 511 YEICPDYLMVMWKRTPTAGDLAFNQCP 538
 RESULT 7
 LITBL HUMAN STANDARD; PRT; 1595 AA.
 AC Q14766;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Latent transforming growth factor beta binding protein, isoform 1L
 DE precursor (ITBP-1) (transforming growth factor beta-1 binding protein
 DE 1) (TGF-beta1-BP-1).
 GN ITBP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-346 FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=96125117; PubMed=8537398;
 RA Olofsson A., Ichijo H., Moren A., ten Dijke P., Miyazono K.,
 Heldin C.-H.;
 RT "Efficient association of an amino-terminally extended form of human
 RT latent transforming growth factor-beta binding protein with the
 RT extracellular matrix.";
 RL J. Biol. Chem. 270:31294-31297 (1995).
 RN [2]
 RP SEQUENCE OF 347-1595 FROM N.A.
 RC TISSUE=Fibroblast, and Platelet;
 RX MEDLINE=90275601; PubMed=2350783;
 RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
 Miyazono K., Claesson-Welsh L., Heldin C.-H.;
 RT "TGF-beta 1 binding protein: a component of the large latent complex
 RT of TGF-beta 1 with multiple repeat sequences.";
 RL Cell 61:1051-1061 (1990).
 RN [3]
 RP INTERACTION WITH FIBRILLIN.
 RX MEDLINE=22421386; PubMed=12429738;
 RA Isogai Z., Ono R.N., Ushiro S., Keene D.R., Chen Y., Mazzieri R.,
 Charbonneau N.L., Reinhardt D.P., Rifkin D.B., Sakai L.Y.;
 RT "Latent transforming growth factor beta-binding protein 1 interacts
 RT with fibrillin and is a microfibril-associated protein.";
 RL J. Biol. Chem. 278:2750-2757 (2003).
 CC -!- SUBUNIT: The large latent complex of TGF-beta1 from platelets is
 CC composed of the TGF-beta1 molecule noncovalently associated with a
 CC disulfide-bonded complex of a dimer of the N-terminal propeptide
 CC of the TGF-beta1 precursor and a third component denoted TGF-
 CC beta1-BP. TGF-beta1-BP does not bind directly to active TGF-BETA1.
 CC Binds to fibrillin.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q14766-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P22064-1; Sequence=External;
 CC -!- TISSUE SPECIFICITY: The long isoform is found in fibroblasts.
 CC -!- PTM: Contains hydroxylated asparagine residues (By similarity).
 CC -!- SIMILARITY: Contains 16 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC -----
 CC EMBL; L48925; AAA96327.1; -;
 CC EMBL; M34057; AAA61160.1; ALT_INIT.
 CC HSSP; P08709; IBF9.
 CC Genew; HGNC:6714; LTBP1.
 CC MIM; 150390; -;
 CC GO; GO:0005578; C:extracellular matrix; NAS.
 CC GO; GO:0005024; F:transforming growth factor-beta receptor ac. .; NAS.
 CC InterPro; IPR000152; Asx_hydroxyl_s.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR002212; Fibril-assoc.
 CC Pfam; PF00008; EGF; 15.
 CC Pfam; PF00683; TB; 4.
 CC SMART; SM00179; EGF_CA; 13.
 CC PROSITE; PS00010; ASX_HYDROXYL; 13.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 11.

[illegible]

FT	DISULFID	1550	1564		BY SIMILARITY.
FT	DISULFID	1566	1579		BY SIMILARITY.
FT	MOD RES	848	848		HYDROXYLATION (BY SIMILARITY).
FT	MOD RES	1011	1011		HYDROXYLATION (BY SIMILARITY).
FT	CARBOHYD	495	495		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1071	1071		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1124	1124		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1240	1240		N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	1595 AA;	173229 MW;	6A091EBA8556D8E5 CRC64;	

Query Match	8.8%;	Score	80.5;	DB	1;	Length	1595;
Best Local Similarity	27.8%;	Pred. NO.	13;				
Matches	40;	Conservative	11;	Mismatches	62;	Indels	31;
						Gaps	5

Qy	14	MVPQWLPRWTWWS	--SSTA-----	WVSSAASALETS	TOPATGATWTKMLH	58
Db	1	MAGAWLRGILLWAGLLASSAHGR	LRRTYVHFGPGAAGALPUSGGP	PASSRTFNVALN	60	
Qy	59	YAGSSRSIPTLEATLTVSPFIASLR	VAR-----	VCLRLC	PYPYKDSSTEPSMRVWPS	112
Db	61	ARYSRSSAAGAPSRASPGVPSERT	TRTSKPGGAALQLRPPPPPPPE	-----	PA	110
Qy	113	CPASLPQALMSSPRWWTCLP	VTK	136		
Db	111	RPAVEGGQLHNP	CGHPAAAPFTK	134		

RESULT	8
YAV2_XANCV	
ID_YAV2_XANCV	STANDARD;
AC	P14728;
DT	01-APR-1990 (Rel. 14, Created)
DT	01-APR-1990 (Rel. 14, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Hypothetical 82 kDa avirulence protein in avrBs3 region.
OS	Xanthomonas campestris (pv. vesicatoria).
OG	Plasmid pXv11.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC	Xanthomonadaceae; Xanthomonas.
OX	NCHI_TaxID=341;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Isolate 71-21;
FX	MEDLINE=89384426; PubMed=2550761;
RT	Bonas U., Stall R.E., Staskawicz B.;
RT	"Genetic and structural characterization of the avirulence gene
RT	avrBs3 from xanthomonas campestris pv. vesicatoria.";
RL	Mol. Gen. Genet. 218:127-136 (1989).
CC	-1- MISCELLANEOUS: THIS IS ONE OF THE HYPOTHETICAL PROTEINS CODED BY
CC	THE OPEN READING FRAMES WITHIN THE REGION REQUIRED FOR AVRBS3
CC	ACTIVITY.

```
CC -----This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X16130; -; NOT ANNOTATED_CDS.
DR PIR; JQ0317; JQ0317.
DR InterPro; IPR005042; Avirulence.
DR Pfam; PF03377; Avirulence; 1.
DR Hypothetical protein; Plasmid; Virulence.
DR KW SEQUENCE 784 AA; 82074 MW; 3B2624B411793744 CRC64;
SQ
Query Match      8.6%; Score 78.5; DB 1; Length 784;
Best Local Similarity 21.7%; Pred. No. 9.7;
Matches 38; Conservative 24; Mismatches 60; Indels 53; Gaps 9;
QY          11 PCWMPV-CWLPHRTWWSSSTAWSVSWASSALETQTQATGATWTKWLHYAGSSRISPTL 69
||::| | : || | : || : || : | : | : | : | : | : | :
```


KW Golgi stack; Alternative splicing
 FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 6 26 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 FT
 FT DOMAIN 27 796 LUMENAL (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1120 1120 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 783 796 SIRVDEEHQOVD -> GSGLCFLAEHPKGG (in isoform Short).
 FT /FTid=VSP 001775.
 FT Missing (in isoform Short).
 FT /FTid=VSP 001776.
 FT
 SQ SEQUENCE 1139 AA; 129282 MW; 01CA3089FDC0028 CRC64;
 Query Match 8.6%; Score 78.5; DB 1; Length 1139;
 Best Local Similarity 21.1%; Pred. No. 14;
 Matches 35; Conservative 23; Mismatches 55; Indels 53; Gaps 7;
 21 WR--TW-----WSSSSTAWVWSSALETSTQPATGATWTK----- 55
 863 WRGCLWTHPHWMTSTSTSWPCTSIQTSTARVQPRRYLKLPLQANFYMPVMAVIQD 922
 56 -----WLHYAGSSRISPLEATLTVSPFLASLRV-----ARVC--LRLLC 93
 923 AQKELTHTAQLGVSSLDQGLEV---ILDRRLQDDNRGLGQGLKDKNKRKTCNRFLL 979
 94 PPYPKDSSTFWSRVPASPASPALQAQLMSSPRWPTCLPVTKLT 139
 980 ERRTVGSEVDHSTSPYSLSLHLSMTYNAP---ALALPVARWQL 1022

RESULT 11
 ARHB RAT STANDARD; PRT; 1527 AA.
 ID ARHB RAT STANDARD; PRT; 1527 AA.
 AC Q9BS67;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Rho guanine nucleotide exchange factor 11 (RhoGEF glutamate transport modulator GTRAP48).
 DE ARHGEF11.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH RHOA; GNA13 AND SLC1A6.
 RX MEDLINE=21133160; PubMed=11242047;
 RA Jackson M., Song W., Liu M.-Y., Jin L., Dykes-Hoberg M., Lin C.-L.G., Bowers W.J., Federoff H.J., Sternweis P.C., Rothstein J.D.;
 RA "Modulation of the neuronal glutamate transporter EAAT4 by two interacting proteins.";
 RL Nature 410:89-93(2001).
 CC -!- FUNCTION: May play a role in the regulation of RhoA GTPase by guanine nucleotide-binding alpha-12 (GNA12) and alpha-13 (GNA13). Acts as guanine nucleotide exchange factor (GEF) for RhoA GTPase and may act as GTPase activating protein (GAP) for GNA12 and GNA13 (By similarity).
 CC -!- SUBUNIT: Interacts with RHOA, GNA13 and SLC1A6. Interacts with GNA12, PLXNB1 and PLXNB2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the membrane upon stimulation (By similarity).
 CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -!- SIMILARITY: Contains 1 RGS-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; AF225961; AAG28597.1; -.
 CC HSSP; Q12923; 3PDZ.
 DR InterPro; IPR001331; GDS CDC24.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000342; Regl Gprotein.
 DR InterPro; IPR000219; RhoGEF.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF00621; RhoGEF; 1.
 DR SMART; SM00228; PDZ; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00315; RGS; 1.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PS00741; DH_1; FALSE_NEG.
 DR PROSITE; PS0010; DH_2; 1.
 DR PROSITE; PS0106; PDZ; 1.
 DR PROSITE; PS00003; PH DOMAIN; 1.
 DR PROSITE; PS0132; RGS; 1.
 KW Guanine-nucleotide releasing factor; GTPase activation; Coiled coil.
 FT DOMAIN 64 143 PDZ.
 FT DOMAIN 323 503 RGS.
 FT DOMAIN 461 487 COILED COIL (POTENTIAL).
 FT DOMAIN 742 931 DH.
 FT DOMAIN 973 1087 PH.
 SQ SEQUENCE 1527 AA; 168533 MW; ABAA20F541A3A9A CRC64;
 Query Match 8.6%; Score 78.5; DB 1; Length 1527;
 Best Local Similarity 28.0%; Pred. No. 19;
 Matches 37; Conservative 11; Mismatches 47; Indels 37; Gaps 5;
 19 LPWETWWSSTAW-----VSWASSALETSTQPA-----TGATWTKWLH 58
 1292 LPPT-----RSSGVNDSPELDNPAAEAASAPSYKVRKVSLLPGGVGAA-----K 1342
 59 YAGSSRISPLEATLTVSPFLASLRVARVCLRLCPYPKDSSTFWSRVPASPASLP 118
 1343 VAGSNAIPDSQSESELSVEGGAQATGNCFYVMPAGPLDSSTPTCTPTSPSQCHSLP 1402
 119 AQLMSSPRWPT 130
 1403 A-----WPT 1406

RESULT 12
 MUSB HUMAN STANDARD; PRT; 5703 AA.
 ID MUSB HUMAN STANDARD; PRT; 5703 AA.
 AC Q9HC84; O00447; O00573; O14985; O15494; O95291; O95451; Q14881;
 AC Q99552; Q9UE28;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High molecular weight salivary mucin MGL) (Sublingual gland mucin).
 GN MUC5B OR MUC5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE OF 1-1594 FROM N.A.
 RA Chen Y., Di Y.P., Wu R.;
 RP "Molecular cloning of the amino-terminal and 5'-flanking region of the human MUC5B mucin gene.";
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE OF 1-1325 FROM N.A.
 RP MEDLINE=99009274; PubMed=9790959;
 RX Offner G.D., Nunes D.P., Keates A.C., Afdhal N.H., Troxler R.F.;

RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL pneumoniae";
CC Nucleic Acids Res. 24:4420-4449(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000046; AAB96113.1; -;
DR PIR; S73791;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 211 AA; 23592 MW; A5E240288852DDB2 CRC64;

Query Match 8.4%; Score 77; DB 1; Length 211;
Best Local Similarity 23.1%; Pred. No. 3.4;
Matches 42; Conservative 12; Mismatches 52; Indels 76; Gaps 11;

QY 15 VPCWLPWRTWMMSSSTAWVSSWASSALETSTQATGATWTKLHYAGSSRISPTL---EA 71
Db 27 VPSWVP-----DPELGAVPKSSALSWTCL-----LLEPRILGALA 63

QY 72 TLTVSFFLASLR-----VARVC--LRLLCPYPKDSSTPSW-----RVAVPS--- 112
Db 64 RLLVSSSIWPLSSSEDFEFTATCNALTLVSPDEP-----HVGWIGIQIOMLKNQWFORPG 118

QY 113 ---CPASLPAQLMSS-----PRWV-----PTCLPVTKLTLRPWNAACGARVK 151
Db 119 VFHCSSRCPPRRSSPSSQTLPRWKYFDRFAAVVSPFPFATAHSTP-----CAARVK 173

QY 152 RR 153
Db 174 RQ 175

RESULT 15
HEP DROME
ID - HEP DROME STANDARD; PRT; 487 AA.
AC Q23577;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dual specificity mitogen-activated protein kinase kinase hemipterous
DE (EC 2.7.1.1-) (MAPKK).
GN HEP OR HEM.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=96067643; PubMed=8521475;
RA Glise B., Bourbon H., Noselli S.;
RT "Hemipterous encodes a novel Drosophila MAP kinase kinase, required
RT for epithelial cell sheet movement.";
RL Cell 83:451-461(1995).
CC -!- FUNCTION: REQUIRED FOR THE EPITHELIAL CELL SHEET MOVEMENT CALLED
CC DORSAL CLOSURE (DC), WHICH ALLOWS ESTABLISHMENT OF THE DORSAL
CC EPIDERMIS. CONTROLS THE EXPRESSION IN THE DORSAL EPITHELIUM EDGES
CC OF ANOTHER DORSAL CLOSURE GENE, PUCKERED (PUC).
CC -!- PTM: MAPKK IS ITSELF DEPENDENT ON SER/THR PHOSPHORYLATION FOR
CC ACTIVITY CATALYZED BY MAP KINASE KINASES (BY SIMILARITY).
CC -!- PTM: WEAKLY AUTOPHOSPHORYLATED
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP
CC kinase kinase subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U05240; AAC46944.1; -;
DR FlyBase; FBgn0010303; hep.
DR GO; GO:0004672; F:protein kinase activity; IDA.
DR GO; GO:0046843; P:dorsal appendage formation; IMP.
DR GO; GO:0003081; P:eggshell pattern formation (sensu Insecta); IMP.
DR GO; GO:0000165; P:MAPKK cascade; NAS.
DR GO; GO:0046844; P:micropyle formation; IMP.
DR GO; GO:0007395; P:spreading of leading edge cells; IMP.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; I.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation; Developmental protein.
FT DOMAIN 95 106 POLY-SER.
FT DOMAIN 195 454 PROTEIN_KINASE.
FT NP_BIND 201 209 ATP (BY SIMILARITY).
FT BINDING 224 224 ATP (BY SIMILARITY).
FT ACT_SITE 318 318 BY SIMILARITY.
FT MOD_RES 346 346 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 350 350 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 487 AA; 53079 MW; 09E248DBD14A1E45 CRC64;

Query Match 8.4%; Score 77; DB 1; Length 487;
Best Local Similarity 30.0%; Pred. No. 8.1;
Matches 33; Conservative 15; Mismatches 38; Indels 24; Gaps 5;

QY 27 SSSSTAWVSWASSALETSTQATGATWTKLHYAGSSRISPTLEATLTVSPFLAS----- 81
Db 99 SSSSSRSRAFRSA-----PATGLRWYTPPTTRVSRATPTL-----PMLSSGPGGD 145

QY 82 LRVAR-VCLRLCLPPVPKDSSTPSRWAVWPCASLPALQMLSSPRWPT 130
Db 146 VECTRPVILPLTPPHPPVSETDMKLKIMEQ-----TGKLNINGRQYPT 190

Search completed: September 13, 2004, 14:19:31
Job time : 29.4737 secs


```

RESULT 2
Q9MIA8      PRELIMINARY;      PRT;      343 AA.
ID Q9MIA8;
AC Q9MIA8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 2.
GN NADH2.
OS Typhlonectes natans (Rubber eel).
OG Typhlonecton.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Gymnophiona; Caeciliidae; Typhlonectes.
OX NCBI_TaxID=8456;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296752; PubMed=10835397;
RA Zardoya R., Meyer A.;
RT "Mitochondrial Evidence on the Phylogenetic Position of Caecilians
  (Amphibia: Gymnophiona).";
RL Genetics 155:765-775(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Zardoya R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154051; AAF78148.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. .; IEA.
DR InterPro; IPR003917; NADH_oxred2.
DR InterPro; IPR001750; Oxidored_g1.
DR Pfam; PF00361; oxidored_g1; 1.
DR PRINTS; PRO1436; NADHDHGNASE2.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
SQ SEQUENCE 343 AA; 37519 MW; 2BAD5B15406EB2C8 CRC64;

Query Match      10.7%; Score 98; DB 8; Length 343;
Best Local Similarity 26.0%; Pred. No. 0.26;
Matches 32; Conservative 13; Mismatches 56; Indels 22; Gaps 4;

QY 28 SSSTAWSSASSA-----LSTSTQATGATWTKW--LHYAGSSRSISPTLEATLTVSPF 78
Db 228 SMTTATWAPATAATLMLVLLSLGLPPTSGPWKMWMLHLEAKQSL-PAIATTAISAL 286

QY 79 LASLRVARVCLRLCCPPYPKDSSTFSPNRVAVPSCPASLPQALMSSPRWPTCLPVTKIT 138
Db 287 LSLFFVRLCYFTTLTQSPNSTNSTPTWRHKQSSTPKLLPATM-----VTALM 334

QY 139 LRP 141
Db 335 LLP 337

RESULT 3
Q96086      PRELIMINARY;      PRT;      332 AA.
ID Q96086;
AC Q96086;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cathepsin L-like tick cysteine proteinase B.
OS Haemaphysalis longicornis.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Haemaphysalis.
OX NCBI_TaxID=44386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=hard tick;
RT "Haemaphysalis longicornis cysteine proteinase gene B.";

```

```

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020491; BAA34704.1; -.
DR HSP; P07711; ICJL.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; peptidase C1.
DR InterPro; IPR000169; SHprot_acfite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PRO0705; PAPA1N.
DR ProDom; PD000158; Peptidase_C1; 2.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Protease; Thiol protease.
SQ SEQUENCE 332 AA; 37001 MW; 8998FCC8D9F0B2FE CRC64;

Query Match      10.6%; Score 97; DB 5; Length 332;
Best Local Similarity 29.5%; Pred. No. 0.31;
Matches 36; Conservative 10; Mismatches 42; Indels 34; Gaps 6;

QY 17 CWLPW--RTWWSSSTAWSSASSALETSTQATGATWTKWLVHYAGSSRSISPTLEATLT 74
Db 158 CWYRWASRTWMTARNPSA-----TTAARAASWTTFSSSTRSRTRASTPRSATPT 205

QY 75 VSPFLASLRVARVCLRLCCPPY-PKDSSTFSPNRVAVPSCPASLPQALMSSPRWPTCLP 133
Db 206 TA-----RTARATSR--PPWGPTAGTLTSPRETRSC-----RRLWFTVGP 246

QY 134 VT 135
Db 247 VS 248

RESULT 4
Q847U5      PRELIMINARY;      PRT;      304 AA.
ID Q847U5;
AC Q847U5;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Aster yellows phytoplasma.
OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
OC Acholeplasmataceae; Phycoplasmata.
OX NCBI_TaxID=35779;
RN [1]
RP SEQUENCE FROM N.A.
RA Melamed S., Tanne E., Ben-Haim R., Edelbaum O., Yogeve D., Sela I.;
RT "A Novel Approach to the Study of the Phytoplasm Genome and the
  Characterization of Sixty Phytoplasmal Genes.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY191277; AAO61941.1; -.
KW Hypothetical protein.
SQ SEQUENCE 304 AA; 34012 MW; 69CF994AB8ADAC2D CRC64;

Query Match      10.5%; Score 96; DB 2; Length 304;
Best Local Similarity 26.1%; Pred. No. 0.35;
Matches 30; Conservative 17; Mismatches 26; Indels 42; Gaps 7;

QY 23 TWWSSTAWVS-W-----ASSALETSTQATGATWTKWLVHYAGSSRSISPTLEATLV 75
Db 17 TRFWSTRSRWTEHPACEKPNARISTVERPASSATW-----CAAISPTTQ----- 63

QY 76 SPFLASLRVARVCLRLCCPPYKD--SSTFSPNRVAVPSCPASLPQALMSSPRWPT 129
Db 64 -----IPNYMWNSSUSPS---SWPACWSSWPTSRSTAP-WTP 97

RESULT 5
Q9GKQ7      PRELIMINARY;      PRT;      834 AA.
ID Q9GKQ7;
AC Q9GKQ7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)

```

```

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative heparan sulfate proteoglycan.
GN NOVOCAN.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Szuchet S., Domowicz M., Arvanitis A., Macklin W.;
RT "The NOVOCans: A Novel Family of Developmentally Regulated
RT Proteoglycans";
RL Mol. Biol. Cell 11:43a-43a(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Szuchet S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF033623; AAD01973.2; -
DR GO: GO:0005515; F:protein binding; IEA.
DR InterPro: IPR000210; BTB_POZ.
DR Pfam: PF00651; BTB; 2.
DR SMART: SM00225; BTB; 2.
DR PROSITE: PS00097; BTB; 1.
SQ SEQUENCE 834 AA; 93070 MW; 756E68FFCB8D8D464 CRC64;

Query Match 10.5%; Score 95.5; DB 6; Length 834;
Best Local Similarity 26.7%; Pred. No. 1.1;
Matches 40; Conservative 13; Mismatches 42; Indels 55; Gaps 11;

QY 14 MVPCLW-----PW-----RTWWSSSTAWVSWASALETSTQ 47
Db 29 VTPWMLSIGDPWLGSPWSPRRRAERWPSWSPTRRAW--WRSSVLEMSST 86
QY 48 ATGATWTKWLHYAGS-SRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSW 106
Db 87 QSGQRSSP---AGTCSGSP--SAMCSGYP-----CRLRLSCPP-----SRRP-- 124
QY 107 RVAMPSCPASLPAQLMSSPRW--PTCLPV 134
Db 125 RSLW--CPTPPAARSALPTSKDPLCADV 152

RESULT 6
ID Q9P4Q1 PRELIMINARY; PRT; 560 AA.
AC Q9P4Q1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Chitinase.
GN CTS1.
OS Ajellomyces capsulata (Histoplasma capsulatum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; Onygenaceae; Ajellomyces.
OX NCBI_TaxID=5037;
RN [1]
RP SEQUENCE FROM N.A.
RA Thomas P.W., Cole G.T.;
RL STRAIN=G217B;
RC "Identification and functional characterization of a chitinase from
RT Histoplasma capsulatum: protein purification and isolation of genomic
RT and full-length cDNA.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF159366; AAF80370.1; -
DR GO: GO:0016798; P:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF001579; Glyco_hydro_18A.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR ProDom: PD000471; Glyco_hydro_18; 1.
DR SMART: SM00636; Glyco_18; 1.

DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 560 AA; 61347 MW; EF34B32ED9B064E2 CRC64;

Query Match 9.9%; Score 90.5; DB 3; Length 560;
Best Local Similarity 26.7%; Pred. No. 2.2;
Matches 40; Conservative 13; Mismatches 56; Indels 41; Gaps 5;

QY 15 VPCWLPWR-TWWSSSTAWVSWASALETSTQPATGATWTKWLHY-----AGSSRISPT 68
Db 51 IDCVSWRSTVMTSTTTITLPPPSASSASASSPPAATSTPAADYPPVISQASESTPPPP 110
QY 69 LEATLVSPFLASLRVARVCLRLCPYPKDS-----TPSWRVAVPSCPASLPAQLMSP 125
Db 111 YEGSCRVRP-----TPKPPSPSPSPSPSPSPSPSPSPSPSPSPSHNMSP 155

QY 126 -----RWPTCLPVTCLT 138
Db 156 DGYKSIVYVNWAIYARNYNQDLFPVKLT 185

RESULT 7
ID Q8CF78 PRELIMINARY; PRT; 321 AA.
AC Q8CF78;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Enolase 1.
GN ENO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK002322; BAC24987.1; -.
DR MGD; MGI:95393; Enol.
DR GO: GO:0000015; C:phosphorylase hydratase complex; IEA.
DR GO: GO:0004634; F:phosphorylase hydratase activity; IEA.
DR GO: GO:0006036; P:glycolysis; IEA.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR Pfam: PF03952; enolase_N; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
SQ SEQUENCE 321 AA; 35108 MW; BFB4D600A4190948 CRC64;

Query Match 9.8%; Score 89.5; DB 11; Length 321;
Best Local Similarity 24.3%; Pred. No. 1.5;
Matches 34; Conservative 13; Mismatches 22; Indels 71; Gaps 8;

QY 1 TVVRLFLAWLPCMVPCWLPWR-TWWSSSTAWVSWASALETSTQPATGATWTKWLHYA 60
Db 236 TLRLSLAWM-----WLP-----PSTGLASMTWTSSLRMT---PAGTSLPTSWL--- 277
QY 61 GSSRISPTLEATLVSPFLASLRVARVCLRLCPYPKDSSTEPSRVAVPSCPASLPAQLM 118
Db 278 -----IC-----TSFSSRTQW--CPSKIPL 296

QY 119 -----AQLMSSPRW 128
Db 297 TRTTGPGSRSLVRASRW 316

RESULT 8

```

```

Q9ENS4
ID Q9ENS4; PRELIMINARY; PRT; 407 AA.
AC Q9ENS4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Thymidine kinase.
GN TK.
OS Human herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
SEQUENCE FROM N.A.
RA Saijo M.;
RT "Thymidine kinase gene of acyclovir-resistant HSV-1.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047366; BAB11943.1; -.
DR HSSP; P03176; 2K15.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004797; F:thymidine kinase activity; IEA.
DR GO; GO:0006230; P:TMP biosynthesis; IEA.
DR InterPro; IPR001889; TK herpes.
DR Pfam; PF00693; TK herpes; 1.
DR Prodom; PD001519; TK herpes; 2.
KW Kinase.
SQ SEQUENCE 407 AA; 44580 MW; 8799AF463A115697 CRC64;

Query Match 9.8%; Score 89.5; DB 12; Length 407;
Best Local Similarity 23.3%; Pred. No. 1.9;
Matches 44; Conservative 11; Mismatches 43; Indels 91; Gaps 11;

QY 16 PCWLPP-----RTWW-----SSSTAWVSWASSALETSQPATGATWKVHLHYAG 61
Db 186 PWRSPSSRRPCPAQTSCWGPFRRTDTSTAW-----PNASAPASGLTLCWPRFAA 237
QY 62 SSRISTLEATLVSPFLASLRVARVCLRLC-----PPYPKDSST 102
Db 238 FTG-----CLPIRCGICRAAGRGGRIGDSFGRPCRP----- 270
QY 103 EPSNRVAVPSCPAS-----LPAQLMSSPWRMP-----TCLPVTKLTLRPPWMAACGARVKR 152
Db 271 VPSFRAIRAHDPISGTRYLPFC--FGPPSCWPPPTATCITCLP-----GPWTSWPNASV-- 320
QY 153 RFLQLTSL 161
Db 321 ---PCTSL 326

RESULT 9
Q9BT89
ID Q9BT89 PRELIMINARY; PRT; 135 AA.
AC Q9BT89;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Similar to RIKEN CDNA 1110017H11 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Straussberg R.;
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC004290; AAH04290.1; -.
SQ SEQUENCE 135 AA; 14299 MW; BA02B4131B64F00C CRC64;

Query Match 9.6%; Score 87.5; DB 4; Length 135;
Best Local Similarity 26.5%; Pred. No. 0.93;
Matches 35; Conservative 8; Mismatches 40; Indels 49; Gaps 9;

QY 20 PWRTWSSSTAWVSWASSALETSQPATGATWKVHLHYAGSSRISPTLEATLVSPFL 79
Db 13 PW--WWMTAETASCVC-----WT-----FTCSACPTNP-S 43
QY 80 ASLRVARVCLRLCPYPKDSSTEP-----SWRVAVPSCPASLPAQLMSSPWRMPVTLPLV 134
Db 44 SGLVITRAC-----CTSSPTATAASSLSGLPMWATAFPFSLGSSTP-----WEPCTP- 89
QY 135 TKLTLRPPWMAAC 146
Db 90 HPLT-RPSGCTC 100

RESULT 10
Q20599
ID Q20599 PRELIMINARY; PRT; 790 AA.
AC Q20599;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
GN F49E10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RL MEDLINE=99069613; PubMed=9851916;
None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA Miller N.;
RL "The sequence of C. elegans cosmid F49E10.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53341; AAC69106.1; -.
DR PIR; T34293; T34293.
DR WormPep; F49E10.2a; CE07263.
DR InterPro; IPR004294; RPE65.
DR Pfam; PF03055; RPE65; 1.
KW Hypothetical protein.
SQ SEQUENCE 790 AA; 88922 MW; B1AE384B8FECB846 CRC64;

Query Match 9.6%; Score 87.5; DB 5; Length 790;
Best Local Similarity 24.8%; Pred. No. 5.9;
Matches 29; Conservative 10; Mismatches 39; Indels 39; Gaps 4;

QY 27 SSSTAWVSWASSALETSQPATGATWKVHLHYAGSSRISPTLEATLVSPFLASLRVAR 86
Db 627 TSTSTTTTTTSTATTTPQTTTTTSEK-----PVTLTQTQTA----- 666
QY 87 VCLRLCPYPKDSSTEPSSRVAVPSCPASLPAQLMSSPWRMPVTLPLRPPW 143
Db 667 -----PPTTVKRTTPQ-----IVPTTTKIPRWNP--LAGSGSTEQPWW 704

RESULT 11
Q95QF5
ID Q95QF5 PRELIMINARY; PRT; 842 AA.
AC Q95QF5;
DT 01-DEC-2001 (Tremblrel. 19, Created)

```


Search completed: September 13, 2004, 14:22:14
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 13:47:14 ; Search time 9.26316 Seconds
(without alignments)
274.521 Million cell updates/sec

Title: US-10-019-219A-2
Perfect score: 63
Sequence: 1 SPRWPTCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	63	100.0	9	4	AAB31702	Aab31702 Peptide f
2	63	100.0	9	5	ABG79087	Abg79087 Human ice
3	63	100.0	162	4	AAB31701	Aab31701 Peptide f
4	63	100.0	166	4	AAB31703	Aab31703 Protein e
5	47	74.6	14	5	ABJ15269	Abj15269 IGF relat
6	46	73.0	72	4	Aau21874	Aau21874 Human car
7	46	73.0	72	7	ADe45842	Ad45842 Human car
8	46	73.0	105	2	Aaw61569	Aaw61569 Vpr prote
9	46	73.0	105	2	Aaw68196	Aaw68196 Vpr bindi
10	46	73.0	256	7	ADb65101	Adb65101 Human pro
11	45	71.4	102	4	AAM06407	Aam06407 Human foe
12	45	71.4	107	4	AAM06855	Aam06855 Human foe
13	44	69.8	136	3	AAg22709	Ag22709 Zea may
14	43	68.3	49	2	AAy25308	Aay25308 HCV NS5B
15	43	68.3	237	4	AAb95763	Aab95763 Human pro
16	43	68.3	238	4	AAU69468	Aau69468 Human pur
17	43	68.3	614	7	ADe6841	Ad6841 Petunia p
18	43	68.3	643	7	ADe6840	Ad6840 Petunia p
19	43	68.3	817	2	AAW48101	Aaw48101 Human dis
20	43	68.3	817	2	AAy30137	Aay30137 Amino aci
21	43	68.3	817	6	ABr82222	Abr82222 Human MBC
22	43	68.3	849	2	AAW48102	Aaw48102 Human dis
23	43	68.3	3010	2	AAW98022	Aaw98022 Infectiou
24	43	68.3	3010	4	AAB59174	Aab59174 Protein e
25	43	68.3	3010	4	AAB31170	Aab31170 Amino aci

ALIGNMENTS

RESULT 1

AAB31702
ID AAB31702 standard; peptide; 9 AA.

XX AAB31702;

XX AC AAB31702;

XX 30-APR-2001 (first entry)

XX Peptide fragment of a human intestinal carboxylesterase (iCE).

XX Intestinal carboxylesterase; iCE; tumour; cytotoxic T lymphocyte; CTL;

XX cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;

XX tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.

XX Homo sapiens.

XX WO200100784-A2.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-FR001791.

XX 28-JUN-1999; 99PR-00008224.

XX (INSR) INST ROUSSY GUSTAVE.

XX Ronsin C, Scott V, Triebel F;

XX WPI; 2001-112443/12.

XX New peptides and its encoding nucleic acid derived from intestinal

XX carboxylesterase, useful as immunostimulants for treating cancer.

XX Claim 2; Page 33; 53pp; French.

XX The present sequence is derived from a human intestinal carboxylesterase
(iCE) polypeptide. iCE induces specific cytotoxic T lymphocytes
(CTL) and secretion by these CTL of cytotoxic factors, e.g. interleukin-
2, interferon gamma and tumour necrosis factor. iCE polypeptides and
polynucleotides are used for treating cancer, by in vivo or in vitro
immunisation, particularly solid cancers and most especially
hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used
to stimulate the immune system, and to increase, in culture, the
production of associated-associated CTL, for reinjection, and/or to
induce secretion of cytotoxic factors from CTL. Dendritic cells loaded
with iCE are used to induce such CTL in cultures

XX Sequence 9 AA;

Query Match 100.0%; Score 63; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPRWPTCL 9
 |||||
 Db 1 SPRWPTCL 9

RESULT 2
 ID ABG79087 standard; peptide; 9 AA.
 XX
 AC ABG79087;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Human ICE class I HLA widely expressed antigen peptide #1.
 XX
 KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen;
 KW lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;
 KW ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell;
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;
 KW cytostatic; human.
 XX
 OS Homo sapiens.
 XX
 FN WO200264057-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 15-FEB-2002; 2002WO-US005212.
 XX
 PR 15-FEB-2001; 2001US-0268687P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Wang R;
 XX
 DR WPI; 2002-627577/67.
 XX
 PT Novel composition for treating a disease in an animal, comprises an
 PT immune effector cell and cell penetrating peptide associated with an
 PT antigen or antibody.
 PS
 PS Disclosure; Page 19; 61pp; English.
 XX
 CC The invention relates to a composition (I) comprising an immune effector
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or
 CC antibody. Also included are (1) a vaccine comprising (I), CPP associated
 CC with an antigen, and a pharmaceutically acceptable carrier and (2)
 CC preparing a composition for a disease, by providing (I) and CPP
 CC associated with an antigen for disease, and introducing the antigen-
 CC associated CPP to (I), where antigen enters into the cell. The antigens
 CC are, for example, tumour antigen derived epitopes recognised by tumour
 CC infiltrating lymphocytes (TIL) of HLA (human leukocyte antigen) class I
 CC or II. The composition is useful for enhancing immunity in an animal to a
 CC disease, by administering a mature dendritic cell comprising CPP
 CC associated with an antigen to disease, to the animal, such that following
 CC the administration, animal is protected from disease, where the animal
 CC comprises both CD4+ and CD8+ T cells. It is also useful for treating a
 CC disease (e.g. cancer, tumour, melanoma, thymoma, lymphoma, sarcoma, lung
 CC cancer, non-Hodgkin's lymphoma, leukaemia, Hodgkin's lymphoma, uterine
 CC cancer, cervical cancer, bladder cancer, kidney cancer, adenocarcinoma,
 CC breast cancer, prostate cancer, ovarian cancer and pancreatic cancer).
 CC The animal is further subjected to a cancer treatment including surgery,
 CC radiation, chemotherapy or gene therapy. The administration of (I),
 CC preferably dendritic cell is prior to, subsequent to or concurrent with,
 CC the cancer treatment. The present sequence is a tumour antigen derived
 CC epitope for inclusion in the composition of the invention

XX Sequence 9 AA;
 SQ
 Query Match 100.0%; Score 63; DB 5; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPRWPTCL 9
 |||||
 Db 1 SPRWPTCL 9

RESULT 3
 AAB31701
 ID AAB31701 standard; peptide; 162 AA.
 XX
 AC AAB31701;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Peptide fragment of a human intestinal carboxylesterase (ICE).
 XX
 KW Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
 KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
 KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
 XX
 OS Homo sapiens.
 XX
 FN WO200100784-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-FR001791.
 XX
 PR 28-JUN-1999; 99FR-00008224.
 XX
 PA (INSR) INST ROUSSY GUSTAVE.
 XX
 PI Ronsin C, Scott V, Triebel F;
 XX
 DR WPI; 2001-112443/12.
 XX
 PT New peptides and its encoding nucleic acid derived from intestinal
 PT carboxylesterase, useful as immunostimulants for treating cancer.
 PS
 PS Claim 1; Page 3; 53pp; French.
 XX
 CC The present sequence is derived from a human intestinal carboxylesterase
 CC (ICE) polypeptide. ICE induces specific cytotoxic T lymphocytes
 CC (CTL) and secretion by these CTL of cytotoxic factors, e.g. interleukin-
 CC 2, interferon gamma and tumour necrosis factor. ICE polypeptides and
 CC polynucleotides are used for treating cancer, by in vivo or in vitro
 CC immunisation, particularly solid cancers and most especially
 CC hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used
 CC to stimulate the immune system, and to increase, in culture, the
 CC production of associated-associated CTL, for reinjection, and/or to
 CC induce secretion of cytotoxic factors from CTL. Dendritic cells loaded
 CC with ICE are used to induce such CTL in cultures
 XX
 SQ Sequence 162 AA;
 Query Match 100.0%; Score 63; DB 4; Length 162;
 Best Local Similarity 100.0%; Pred. No. 0.045;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPRWPTCL 9
 |||||
 Db 124 SPRWPTCL 132

RESULT 4
 AAB31703
 ID AAB31703 standard; protein; 166 AA.

```

XX AC AAB31703;
XX DT 30-APR-2001 (first entry)
XX DE Protein encoded by an intestinal carboxylesterase (iCE) cDNA.
XX KW Intestinal carboxylesterase; iCE; tumour; cytotoxic T lymphocyte; CTL;
XX KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
XX KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney.
XX OS Homo sapiens.
XX PN WO200100784-A2.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-FR001791.
XX PR 28-JUN-1999; 99FR-00008224.
XX PA (INSR ) INST ROUSSY GUSTAVE.
XX PI Ronnin C, Scott V, Triebel F;
XX DR WPI; 2001-112443/12.
XX DR N-FSDB; AAF25258.
XX PT New peptides and its encoding nucleic acid derived from intestinal
XX PT carboxylesterase, useful as immunostimulants for treating cancer.
XX PS Disclosure; Fig 8A; 53pp; French.
XX CC The present sequence is encoded by the coding region of human intestinal
XX CC carboxylesterase (iCE) gene. iCE induces specific cytotoxic T
XX CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors, e.g.
XX CC interleukin-2, interferon gamma and tumour necrosis factor. iCE
XX CC polypeptides and polynucleotides are used for treating cancer, by in vivo
XX CC or in vitro immunisation, particularly solid cancers and most especially
XX CC hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used
XX CC to stimulate the immune system, and to increase, in culture, the
XX CC production of associated-associated CTL, for reinjection, and/or to
XX CC induce secretion of cytotoxic factors from CTL. Dendritic cells loaded
XX CC with iCE are used to induce such CTL in cultures.
XX SQ Sequence 166 AA;
XX Query Match 100.0%; Score 63; DB 4; Length 166;
XX Best Local Similarity 100.0%; Pred. No. 0.046;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 1 SPRWPTCL 9
XX Db 128 SPRWPTCL 136
XX RESULT 5
XX ABJ15269
XX ID ABJ15269 standard; peptide; 14 AA.
XX AC ABJ15269;
XX XX
XX DT 16-JAN-2003 (first entry)
XX DE IGF related native phage peptide SEQ ID No 45.
XX KW Cytostatic; antidiabetic; osteopathic; vasotropic; tranquiliser; IGF-1;
XX KW vulnerary; antiasthmatic; ophthalmological; antagonise; ischemic injury;
XX KW insulin-like growth hormone 1; IGF; cancer; diabetic; nephropathy;
XX KW diabetic retinopathy; acromegaly; macular degeneration; trauma; asthma;
XX KW restenosis.
XX OS Unidentified.

```

```

XX PN WO200272780-A2.
XX PD 19-SEP-2002.
XX PF 13-MAR-2002; 2002WO-US007606.
XX PR 14-MAR-2001; 2001US-0275904P.
XX PA (GETH ) GENENTECH INC.
XX XX
XX PI Deshayes K, Lowman HB, Schaffer ML, Sidhu SS;
XX DR WPI; 2002-732826/79.
XX CC New peptides antagonizing insulin-like growth factor (IGF), useful for
XX CC treating disorder such as cancer, diabetic complication exacerbated by
XX CC IGF-1, acromegaly, age-related macular degeneration, ischemic injury,
XX CC trauma, asthma.
XX PS Example 1; Page 42; 86pp; English.
XX CC The invention relates to novel peptides that can antagonise the
XX CC interaction of insulin-like growth hormone 1 (IGF-1). The peptides are
XX CC useful for treating disorders such as cancer, diabetic complication
XX CC exacerbated by IGF-1, e.g. diabetic retinopathy or nephropathy,
XX CC acromegaly, age-related macular degeneration, ischemic injury or trauma.
XX CC Other disorders that can be treated by the peptide include restenosis or
XX CC asthma. This sequence represents a peptide relating to the IGF antagonist
XX CC peptides of the invention
XX SQ Sequence 14 AA;
XX Query Match 74.6%; Score 47; DB 5; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 4 WWPCTL 9
XX Db 8 WWPCTL 13
XX RESULT 6
XX AAU21874
XX ID AAU21874 standard; protein; 72 AA.
XX AC AAU21874;
XX DT 17-DEC-2001 (first entry)
XX DE Human cardiovascular system antigen polypeptide SEQ ID No 648.
XX KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
XX KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
XX KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
XX KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
XX KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
XX KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
XX KW cerebrovascular disorder; nervous system disorder; bacterial infection;
XX KW fungal infection; viral infection; ocular disorder; endocrine disorder;
XX KW gastrointestinal disorder; renal disorder; respiratory disorder;
XX KW wound healing; skin aging; organ transplantation; tissue regeneration;
XX KW anti-infertility.
XX OS Homo sapiens.
XX PN WO200155321-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001340.
XX PR 31-JAN-2000; 2000US-0179065P.

```

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184564P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-MAR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-451930/48.
DR N-PSDB; AAS35148.
XX
XX New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system.

XX Claim 11; SEQ ID NO 648; 674pp; English.

XX Sequences AAU21852-AAU22466 represent the cardiovascular system antigen

CC polypeptides of the invention. Cardiovascular system antigens and their

CC associated polynucleotides are useful in the diagnosis, treatment and

CC prevention of various types of disorders in e.g. humans, mice, rabbits,

CC goats, horses, cats, dogs, chickens or sheep. A pathological condition

CC can be determined by detecting the presence or absence of a mutation in a

CC cardiovascular system antigen polynucleotide. The treatable disorders

CC include autoimmune diseases such as rheumatoid arthritis,

CC hyperproliferative disorders such as neoplasms of the breast or liver,

CC cardiovascular disorders such as cardiac arrest, cerebrovascular

CC disorders such as cerebral ischaemia, nervous system disorders such as

CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,

CC ocular disorders such as corneal infection, endocrine disorders such as

CC premature labour and infertility, gastrointestinal disorders such as

CC Crohn's disease, renal disorders such as glomerulonephritis and

CC respiratory disorders such as asthma and pleurisy. The polypeptides can

CC also be used to aid wound healing, to prevent skin aging due to sunburn,

CC to maintain organs before transplantation, to regenerate tissues and in

CC chemotaxis. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

Query Match 73.0%; Score 46; DB 4; Length 72;
Best Local Similarity 85.7%; Pred. No. 8;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWWPTCL 9
|||||

Db 49 RWWPPCL 55

RESULT 7

ID ADE45842 standard; protein; 72 AA.

XX AC ADE45842;

XX DT 29-JAN-2004 (first entry)

XX DE Human cardiovascular system related polypeptide #23.

XX Human; cardiovascular system related polypeptide; cancer;
KW proliferative disorder; foetal abnormality; developmental abnormality;
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
endocrine disorder.

XX Homo sapiens.

XX US2003059908-A1.

XX PD 27-MAR-2003.

XX PF 07-MAR-2002; 2002US-00091504.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-02411787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-02519097P.
 PR 05-JAN-2001; 2000US-0254097P.
 PR 17-JAN-2001; 2000US-0259678P.
 PR 17-JAN-2001; 2000US-00764869.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 XX WPI: 2003-743766/70.
 DR N-PSDB; ADE45227.
 XX
 XX New cardiovascular system related polynucleotides and polypeptides,
 PT useful for preventing, treating, or ameliorating a medical condition,
 PT such as cancer of cardiovascular tissues and cancer metastases.
 XX
 XX Claim 11; SEQ ID NO 648; 262pp; English.
 PS
 CC The invention relates to human cardiovascular system related polypeptides
 CC and the polynucleotides encoding them. The polypeptides, polynucleotides
 CC and antibodies to the polypeptides are useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition,
 CC for preventing, treating, or ameliorating a medical condition, such as

CC cancer of cardiovascular system tissues, proliferative disorders, foetal
 CC and developmental abnormalities, haematopoietic disorders, diseases of
 CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
 CC arthritis), inflammation, allergies, neurological disorders (e.g.,
 CC Alzheimer's disease, Parkinson's disease), cognitive disorders,
 CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,
 CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic
 CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
 CC related disorders, endocrine disorders and infections. The nucleic acids
 CC are also useful for chromosome identification, radiation hybrid mapping
 CC or long-range restriction mapping. The polypeptides and polynucleotides
 CC may also be used as food additives or preservatives to increase or
 CC decrease storage capabilities, fat content or other nutritional
 CC components. This sequence represents a human cardiovascular system
 CC related polypeptide of the invention.
 XX
 SQ Sequence 72 AA;
 Query Match 73.0%; Score 46; DB 7; Length 72;
 Best Local Similarity 85.7%; Pred. No. 8;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 RWWPTCL 9
 DB 49 RWWPPCL 55
 |||||
 RESULT 8
 AAW61569
 ID AAW61569 standard; peptide; 105 AA.
 XX
 AC AAW61569;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Vpr protein binding B29-1 amino acid sequence.
 XX
 KW Lentiviral infection; Vpr protein; HIV infection; cell stasis;
 KW cell death.
 XX
 OS Homo sapiens.
 XX
 PN WO9835234-A1.
 XX
 PD 13-AUG-1998.
 XX
 PF 11-FEB-1998; 98WO-US003008.
 XX
 PR 11-FEB-1997; 97US-00797907.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Chen ISY, Jowett JEM, Withers-Ward E;
 XX
 DR WPI: 1998-447375/38.
 XX
 PT Identification of compounds binding the HIV-1 Vpr protein - that block
 PT Vpr-mediated cell stasis, useful for treating HIV-1 or other lentiviral
 PT infections.
 XX
 PS Disclosure; Fig 8; 63pp; English.
 XX
 CC This represents the amino acid sequence of the B29-1 protein. This
 CC protein can bind to the Vpr protein encoded by the HIV genome. The
 CC invention provides a method of identifying an agent for use in treating
 CC lentiviral infections. The method comprises contacting a cellular target
 CC of the Vpr protein with the agent to be tested, and assessing the ability
 CC of the agent to block interaction of the Vpr protein with the cellular
 CC target, where an agent which blocks this interaction is an anti-
 CC lentiviral agent. Alternatively, the agent contacts a cell expressing the
 CC Vpr protein under conditions where the Vpr protein induces cell stasis in
 CC the absence of the agent. Identification of the agent is then observed by
 CC blockage of Vpr-induced cell stasis. The method allows the identification

CC of compounds that block Vpr-mediated cell stasis and ultimately cell
 CC death. The compounds can thus be used in the treatment of HIV and other
 CC lentiviral infections

XX SQ Sequence 105 AA;
 Query Match 73.0%; Score 46; DB 2; Length 105;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SPRWPTC 8
 | |||| |
 60 SSRWPAC 67

Db

RESULT 9
 AAW68196
 ID AAW68196 standard; peptide; 105 AA.
 XX AC AAW68196;
 XX AC AAW68196;
 XX AC AAW68196;
 DT 25-MAR-2003 (revised)
 DT 29-OCT-1998 (first entry)
 XX AC AAW68196;
 DE Vpr binding protein B29-1 amino acid sequence.
 XX AC AAW68196;
 KW LentiVirus; Vpr protein; HIV infection; cell stasis; cell death; cancer;
 KW auto-immune disease; B29-1.
 XX OS Homo sapiens.
 XX WO9835032-A2.
 XX 13-AUG-1998.
 XX 11-FEB-1998; 98WO-US003390.
 XX 11-FEB-1997; 97US-00798597.
 XX 24-OCT-1997; 97US-00959279.
 XX (REGC) UNIV CALIFORNIA.
 XX Chen ISY, Jowett JBM, Withers-Ward E, Stewart SA, Poon B;
 PI Feigon J, Dieckmann T;
 XX WPI; 1998-447229/38.
 XX Arresting cell growth using lentivirus Vpr virion protein - used for
 PT treatment of cancer and screening for agents that reduce Vpr binding,
 PT e.g. anti-HIV agents.
 XX Disclosure; Fig 7; 7lpp; English.
 XX This represents the amino acid sequence of the B29-1 protein. This
 CC protein can bind to the Vpr protein encoded by the HIV genome. This is
 CC used as a cellular target in the method of the invention of identifying
 CC antitumour therapeutic candidates. The invention provides a method for
 CC arresting the growth of a cell by treatment with a Vpr lentivirus protein
 CC or its analogue. Agents that reduce binding of Vpr to a cellular target
 CC are useful for treating HIV (human immune deficiency virus) infection or
 CC more generally for restoring growth. The antitumour agent identified is
 CC useful for treating any type of cancer, since it induces cell stasis
 CC (blocks development at the G2 stage) and death. The agents can also be
 CC used for treating autoimmune diseases. (Updated on 25-MAR-2003 to correct
 CC PI field.)

XX SQ Sequence 105 AA;
 Query Match 73.0%; Score 46; DB 2; Length 105;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SPRWPTC 8

Db 60 SSRWPAC 67

RESULT 10

ADB65101
ID ADB65101 standard; protein; 256 AA.

XX AC ADB65101;

XX DT 04-DEC-2003 (first entry)

XX DE Human protein encoded by clone SPLEN20015100.

XX KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.

XX OS Homo sapiens.

XX FN EP1308459-A2.

XX AC ADB65101;

XX DT 07-MAY-2003.

XX PF 28-MAR-2002; 2002EP-00007401.

XX PR 05-NOV-2001; 2001JP-00379298.

XX PR 25-JAN-2002; 2002US-00350978.

XX PA (HELI-) HELIX RES INST.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX DR WPI; 2003-450961/43.

XX DR N-PSDB; ADB63131.

XX PT New polynucleotides and polypeptides, useful for developing a diagnostic

XX PT marker or medicines for regulation of their expression and activity, or

XX PT as targets of gene therapy.

XX PS Claim 1; Page; 22pp; English.

XX CC The invention discloses a polynucleotide comprising a sequence selected

XX CC from 1970 fully defined nucleotide sequences which encode novel

XX CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide

XX CC or its partial peptide, an antibody binding to the polypeptide or peptide

XX CC of the polynucleotide, immunologically assaying the polypeptide or

XX CC peptide of the polynucleotide by contacting the polypeptide or peptide

XX CC with the antibody of the encoded protein, and observing the binding

XX CC between the two, a transformant carrying the polynucleotide in an

XX CC expressible manner and an antisense polynucleotide. The oligonucleotide

XX CC is useful as a primer for synthesising the polynucleotide, or as a probe

XX CC for detecting the polynucleotide. The polynucleotides and encoded

XX CC proteins are useful as pharmaceutical agents and many disease-related

XX CC genes may be included in them, for developing a diagnostic marker or

XX CC medicines for regulation of their expression and activity, or as targets

XX CC of gene therapy. The genes are involved in tissue and/or cell

XX CC regeneration. Membrane proteins, signal transduction-related proteins,

XX CC transcription-related proteins, disease-related proteins and genes

XX CC encoding them can be used as indicators for diseases (e.g. osteoporosis,

XX CC neurological diseases, cancer, tumours. The cDNA may be used to regulate

XX CC the activity or expression of the encoded protein to treat diseases. The

XX CC sequence presented is a protein of the invention. Note: Some of the

XX CC sequence data for this patent is not represented in the printed

XX CC specification, but is based on sequence information supplied by the

XX CC European Patent Office.

XX SQ Sequence 256 AA;

Query Match 73.0%; Score 46; DB 7; Length 256;
 Best Local Similarity 75.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTC 8
 | | | | |
 Db 211 SSRWWPAC 218

RESULT 11
 AAM06407
 ID AAM06407 standard; protein; 102 AA.
 XX
 AC AAM06407;
 XX
 DT 05-OCT-2001 (first entry)
 XX
 DE Human foetal protein, SEQ ID NO: 138.
 XX
 KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nontropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200153339-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002723.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 15-SEP-2000; 2000US-00663870.
 PR 06-NOV-2000; 2000US-00707351.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werthman T;
 XX
 DR WPI; 2001-465571/50.
 DR N-PSDB; AAH94082.
 XX
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 PT associated with dysfunction of the protein e.g. cancers, immune
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation.
 XX
 PS Claim 10; Page 231; 715pp; English.
 XX
 CC The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are useful in
 CC the treatment and diagnosis of diseases such as cancers, immune
 CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
 CC nervous system disorders and inflammation. The present sequence is a
 CC polypeptide encoded by a cDNA assembled using an expressed sequence tag
 CC (EST) found to be expressed in human foetal tissue cDNA libraries

Query Match 71.4%; Score 45; DB 4; Length 102;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTC 8
 | | | | |
 Db 16 SSRWWPVC 23

RESULT 12
 AAG22709
 ID AAG22709 standard; protein; 136 AA.
 XX
 AC AAG22709;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Zea mays protein fragment SEQ ID NO: 25743.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;

AAM06855
 ID AAM06855 standard; protein; 107 AA.
 XX
 AC AAM06855;
 XX
 DT 05-OCT-2001 (first entry)
 XX
 DE Human foetal protein, SEQ ID NO: 1063.
 XX
 KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nontropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200153339-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002723.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 15-SEP-2000; 2000US-00663870.
 PR 06-NOV-2000; 2000US-00707351.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werthman T;
 XX
 DR WPI; 2001-465571/50.
 DR N-PSDB; AAH94530.
 XX
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 PT associated with dysfunction of the protein e.g. cancers, immune
 PT disorders, growth disorders, thrombolytic disorders, nervous system
 PT disorders and inflammation.
 XX
 PS Example 4; Page 600-601; 715pp; English.
 XX
 CC The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are useful in
 CC the treatment and diagnosis of diseases such as cancers, immune
 CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
 CC nervous system disorders and inflammation. The present sequence is a
 CC polypeptide encoded by a cDNA assembled using an expressed sequence tag
 CC (EST) found to be expressed in human foetal tissue cDNA libraries

Query Match 71.4%; Score 45; DB 4; Length 107;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTC 8
 | | | | |
 Db 16 SSRWWPVC 23

RESULT 13
 AAG22709
 ID AAG22709 standard; protein; 136 AA.
 XX
 AC AAG22709;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Zea mays protein fragment SEQ ID NO: 25743.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139117P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 69.8%; Score 44; DB 3; Length 136;
 Best Local Similarity 85.7%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTC 8
 DB 24 PRWWWTC 30

RESULT 14
 AAY25308
 ID AAY25308 standard; protein; 49 AA.

XX AC AAY25308;

XX DT 03-SEP-1999 (first entry)

XX DE HCV NS5B carboxy-terminus protein fragment from genotype HCV J491.

XX NS5B; antiviral compound; truncated protein; mutant; immunoprotective;
 KW vaccine; immunological response; protection; disease; Flaviviridae;
 KW virus; x-ray crystallography; hydrophobic tail.

XX OS Hepatitis C virus.

XX PN WO9929843-A1.

XX

PD 17-JUN-1999.

XX PF 09-DEC-1998; 98WO-US026070.

XX PR 11-DEC-1997; 97US-0069208P.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PI Del Vecchio A;

XX WPI; 1999-404939/34.

XX Nucleic acid encoding Hepatitis C Virus NS5B truncated protein.

XX Disclosure; Page 16; 63pp; English.

XX This invention describes a novel nucleic acid encoding a Hepatitis C
 CC Virus (HCV) truncation mutant of NS5B. The product of the invention has
 CC immunoprotective activity and can be used as a vaccine. The NS5B mutant
 CC or a nucleic acid vector directing expression of the mutant, are used to
 CC induce an immunological response in a mammal. Antibodies against the NS5B
 CC mutant are used to protect mammals against diseases caused by viruses of
 CC the Flaviviridae. Additionally, soluble protein produced by the method
 CC would allow for determination of the structure of the protein via x-ray
 CC crystallography or other known methods. The HCV NS5B truncation mutant
 CC has a deletion of a hydrophobic tail which release the protein into the
 CC soluble portion of the cell, allowing for a greater recovery of soluble
 CC protein for screening for inhibitors of NS5B enzymatic activity. AAY25293
 CC -Y25322 represent the carboxy-terminus of the NS5B protein from various
 CC HCV genotypes which are used to describe the method of the invention

SQ Sequence 49 AA;

Query Match 68.3%; Score 43; DB 2; Length 49;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTCL 9
 DB 27 PRWWFLCL 34

RESULT 15

AAB95763
 ID AAB95763 standard; protein; 237 AA.

XX AC AAB95763;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:18691.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX

DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 18691; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 237 AA;

Query Match 68.3%; Score 43; DB 4; Length 237;
 Best Local Similarity 100.0%; Pred.No. 70;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 WWPTC 8
 Db 99 WWPTC 103

Search completed: September 13, 2004, 14:19:08
 Job time : 18.2632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:16:17 : Search time 2.36842 Seconds
(without alignments)
196.179 Million cell updates/sec

Title: US-10-019-219A-2
Perfect score: 63
Sequence: 1 SPRWWPTCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCtus_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	77.8	222	4	US-09-252-991A-32998
2	43	68.3	144	3	US-09-208-140-20
3	43	68.3	144	4	US-09-252-991A-20518
4	43	68.3	3010	3	US-09-014-416-3
5	42	66.7	82	4	US-09-252-991A-18536
6	42	66.7	148	4	US-09-252-991A-28538
7	42	66.7	171	4	US-09-252-991A-16829
8	41	65.1	156	4	US-09-252-991A-23897
9	41	65.1	345	4	US-09-134-000C-5047
10	40	63.5	98	4	US-09-252-991A-26215
11	40	63.5	125	4	US-09-252-991A-19066
12	40	63.5	194	4	US-09-252-991A-24154
13	40	63.5	360	4	US-09-252-991A-23898
14	40	63.5	520	1	US-08-261-822A-10
15	40	63.5	520	5	PCT-US95-07744A-10
16	40	63.5	632	4	US-09-252-991A-23129
17	39.5	62.7	227	4	US-09-489-039A-7970
18	39.5	62.7	351	4	US-09-252-991A-17990
19	39	61.9	144	4	US-09-252-991A-17313
20	39	61.9	169	4	US-09-252-991A-32019
21	39	61.9	239	4	US-09-252-991A-30037
22	39	61.9	246	4	US-09-252-991A-25102
23	39	61.9	339	4	US-09-134-000C-4881
24	39	61.9	376	2	US-08-758-621-10
25	39	61.9	376	3	US-09-107-858-10
26	39	61.9	376	4	US-09-579-174-10
27	39	61.9	458	4	US-09-252-991A-30535

28	39	61.9	552	4	US-09-252-991A-22692	Sequence 22692, A
29	39	61.9	567	1	US-08-261-822A-12	Sequence 12, Appl
30	39	61.9	567	5	PCT-US95-07744A-12	Sequence 12, Appl
31	38.5	61.1	117	4	US-09-252-991A-29121	Sequence 29121, A
32	38.5	61.1	224	4	US-09-252-991A-21129	Sequence 21129, A
33	38.5	61.1	358	4	US-09-252-991A-16785	Sequence 16785, A
34	38	60.3	6	4	US-09-007-288E-39	Sequence 39, Appl
35	38	60.3	8	4	US-09-007-288E-73	Sequence 73, Appl
36	38	60.3	8	4	US-09-007-288E-75	Sequence 75, Appl
37	38	60.3	8	4	US-09-007-288E-75	Sequence 75, Appl
38	38	60.3	106	4	US-09-252-991A-21103	Sequence 21103, A
39	38	60.3	246	4	US-09-252-991A-24028	Sequence 24028, A
40	38	60.3	365	4	US-09-252-991A-31971	Sequence 31971, A
41	38	60.3	370	4	US-09-252-991A-27810	Sequence 27810, A
42	38	60.3	390	4	US-09-252-991A-20713	Sequence 20713, A
43	38	60.3	435	4	US-09-491-577-54	Sequence 54, Appl
44	38	60.3	584	1	US-08-261-822A-8	Sequence 8, Appl
45	38	60.3	584	5	PCT-US95-07744A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-32998
; Sequence 32998, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32998
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32998

Query Match 77.8%; Score 49; DB 4; Length 222;
Best Local Similarity 87.5%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 SPRWWPTC 8
Db 190 SCRWWPTC 197

RESULT 2
US-09-208-140-20
; Sequence 20, Application US/09208140
; Patent No. 6228576
; GENERAL INFORMATION:
; APPLICANT: Del Vecchio, Alfred
; TITLE OF INVENTION: HEPATITIS C VIRUS NS5B TRUNCATED PROTEIN
; TITLE OF INVENTION: AND METHODS THEREOF TO IDENTIFY ANTIVIRAL COMPOUNDS
; FILE REFERENCE: P50743
; CURRENT APPLICATION NUMBER: US/09/208,140
; CURRENT FILING DATE: 1998-12-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-09-208-140-20

Query Match 68.3%; Score 43; DB 3; Length 49;
 Best Local Similarity 75.0%; Pred. No. 3.5;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTCL 9
 |||:|
 DB 27 PRWFFLCL 34

RESULT 3
 US-09-252-991A-20518
 ; Sequence 20518, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20518
 ; LENGTH: 144
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20518

Query Match 68.3%; Score 43; DB 4; Length 144;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWPCTC 8
 |||
 DB 70 WWPCTC 74

RESULT 4
 US-09-014-416-3
 ; Sequence 3, Application US/09014416
 ; Patent No. 6153421
 ; GENERAL INFORMATION:
 ; APPLICANT: Yanagi, Masayuki
 ; APPLICANT: Bukh, Jens
 ; APPLICANT: Emerson, Susanne U.
 ; APPLICANT: Purcell, Robert H.
 ; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
 ; TITLE OF INVENTION: US95 THEREOF
 ; FILE REFERENCE: 20264276
 ; CURRENT APPLICATION NUMBER: US/09/014,416
 ; CURRENT FILING DATE: 1998-01-27
 ; EARLIER APPLICATION NUMBER: US 60/053,062
 ; EARLIER FILING DATE: 1997-07-18
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 3010
 ; TYPE: PRT
 ; ORGANISM: Hepatitis C virus
 US-09-014-416-3

Query Match 68.3%; Score 43; DB 3; Length 3010;
 Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTCL 9
 |||:|
 DB 2988 PRWFFLCL 2995

RESULT 5
 US-09-252-991A-18536
 ; Sequence 18536, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 18536
 ; LENGTH: 82
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-18536

Query Match 66.7%; Score 42; DB 4; Length 82;
 Best Local Similarity 85.7%; Pred. No. 8.3;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTC 8
 |||
 DB 30 PRWWPTC 36

RESULT 6
 US-09-252-991A-28538
 ; Sequence 28538, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 28538
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-28538

Query Match 66.7%; Score 42; DB 4; Length 148;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWWPTC 8
 :|||
 DB 7 TPRWSPCT 14

RESULT 7
 US-09-252-991A-16829
 ; Sequence 16829, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18

; Sequence 24154, Application US/09252991A

```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24154
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24154

Query Match 63.5%; Score 40; DB 4; Length 194;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPT 7
Db 60 PSWPT 65

RESULT 13
US-09-252-991A-23898
; Sequence 23898, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23898
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-23898

Query Match 63.5%; Score 40; DB 4; Length 360;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWPTC 8
Db 31 SPRSWPAC 38

RESULT 14
US-08-261-822A-10
; Sequence 10, Application US/08261822A
; Patent No. 5650553
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23898
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-23898

Query Match 63.5%; Score 40; DB 4; Length 360;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWPTC 8
Db 31 SPRSWPAC 38

RESULT 14
US-08-261-822A-10
; Sequence 10, Application US/08261822A
; Patent No. 5650553
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA

```

```

; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-261-822A-10

Query Match 63.5%; Score 40; DB 1; Length 520;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWPT 7
Db 216 TPPWPT 222

RESULT 15
PCT-US95-07744A-10
; Sequence 10, Application PC/TUS9507744A
; GENERAL INFORMATION:
; APPLICANT: Trustees of The University of Pennsylvania
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07744A
; FILING DATE: 15-JUNE-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,822
; FILING DATE: June 17, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids

```

us-10-019-219a-2.ra1

Wed Sep 15 10:54:27 2004

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PCT-US95-07744A-10

Query Match 63.5%; Score 40; DB 5; Length 520;
Best Local Similarity 71.4%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWPT 7
: | |||
Db 216 TPFWPT 222

Search completed: September 13, 2004, 14:23:55
Job time : 2.36842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:22:27 ; Search time 9.21053 Seconds
(without alignments)
313.360 Million cell updates/sec

Title: US-10-019-219A-2
Perfect score: 63
Sequence: 1 SPRWWPTCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	9	16	US-10-447-161-96
2	47	74.6	14	14	US-10-098-093-45
3	47	74.6	70	16	US-10-437-963-167296
4	47	74.6	181	14	US-10-029-386-31530
5	46	73.0	34	14	US-10-029-386-31685
6	46	73.0	72	9	US-09-764-869-648
7	46	73.0	72	14	US-10-091-504-648
8	46	73.0	72	15	US-10-227-577-648
9	45	73.0	256	15	US-10-104-047-3255
10	45	71.4	371	16	US-10-437-963-152726
11	44	69.8	102	16	US-10-437-963-159523
12	44	69.8	248	15	US-10-394-575-81
13	43	68.3	49	9	US-09-733-183A-20
14	43	68.3	49	14	US-10-342-372-20
15	43	68.3	237	12	US-10-112-944-429

16	43	68.3	238	14	US-10-204-887-116	Sequence 116, App
17	42.5	67.5	309	15	US-10-374-780A-609	Sequence 609, App
18	42.5	67.5	309	16	US-10-437-963-146364	Sequence 146364, A
19	42	66.7	34	9	US-09-864-761-42036	Sequence 42036, A
20	42	66.7	121	12	US-10-424-599-151028	Sequence 151028, A
21	42	66.7	153	15	US-10-264-049-3103	Sequence 3103, App
22	42	66.7	264	12	US-10-425-114-43726	Sequence 43726, A
23	42	66.7	651	12	US-10-152-886-19	Sequence 19, Appl
24	41.5	65.9	104	12	US-10-424-599-185801	Sequence 185801, A
25	41	65.1	51	12	US-10-424-599-194164	Sequence 194164, A
26	41	65.1	68	10	US-09-986-480-449	Sequence 449, App
27	41	65.1	81	12	US-10-425-114-49774	Sequence 49774, A
28	41	65.1	84	16	US-10-437-963-156184	Sequence 156184, A
29	41	65.1	92	12	US-10-424-599-247576	Sequence 247576, A
30	41	65.1	152	14	US-10-156-761-14346	Sequence 14346, A
31	41	65.1	187	16	US-10-767-701-51785	Sequence 51785, A
32	41	65.1	240	12	US-10-425-114-60647	Sequence 60647, A
33	41	65.1	300	16	US-10-437-963-157722	Sequence 157722, A
34	41	65.1	629	16	US-10-437-963-204503	Sequence 204503, A
35	40	63.5	59	12	US-10-424-599-220251	Sequence 220251, A
36	40	63.5	147	12	US-10-424-599-165824	Sequence 165824, A
37	40	63.5	155	9	US-09-729-674-104	Sequence 104, App
38	40	63.5	155	16	US-10-437-963-176142	Sequence 176142, A
39	40	63.5	166	12	US-10-424-599-265903	Sequence 265903, A
40	40	63.5	444	16	US-10-437-963-180946	Sequence 180946, A
41	40	63.5	493	16	US-10-437-963-180945	Sequence 180945, A
42	40	63.5	518	10	US-09-934-455-148	Sequence 148, App
43	40	63.5	518	12	US-10-225-066A-570	Sequence 570, App
44	40	63.5	518	15	US-10-374-780A-2466	Sequence 2466, App
45	40	63.5	622	12	US-10-425-114-54091	Sequence 54091, A

ALIGNMENTS

RESULT 1
US-10-447-161-96
; Sequence 96, Application US/10447161
; Publication No. US20040023314A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484US1
; CURRENT APPLICATION NUMBER: US/10/447,161
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-447-161-96

Query Match 100.0%; Score 63; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPRWWPTCL 9
| | | | |
Db 1 SPRWWPTCL 9

RESULT 2
US-10-098-093-45
; Sequence 45, Application US/10098093
; Publication No. US20030092631A1
; GENERAL INFORMATION:
; APPLICANT: Deshayes, Kurt D.
; APPLICANT: Lowman, Henry B.

; APPLICANT: Schaffer, Michelle L.
 ; APPLICANT: Sidhu, Sachdev S.
 ; TITLE OF INVENTION: IGF ANTAGONIST PEPTIDES
 ; FILE REFERENCE: P1863R1
 ; CURRENT APPLICATION NUMBER: US/10/098,093
 ; CURRENT FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: US 60/275,904
 ; PRIOR FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 122
 ; SEQ ID NO 45
 ; LENGTH: 14
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Sequence is synthesized
 US-10-098-093-45

Query Match 74.6%; Score 47; DB 14; Length 14;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWPTCL 9
 |||||
 Db 8 WWPTCL 13

RESULT 3
 US-10-437-963-167296
 ; Sequence 167296, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 167296
 ; LENGTH: 70
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(70)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_65920C.1.pap
 US-10-437-963-167296

Query Match 74.6%; Score 47; DB 16; Length 70;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 2 PRWNP---TCL 9
 |||||
 Db 5 PRWPDGIFTCL 16

RESULT 4
 US-10-029-386-31530
 ; Sequence 31530, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 31530
 ; LENGTH: 181
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AF196968.1
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.8
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 49
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
 ; OTHER INFORMATION: SWISSPROT HIT: P25304, EVALUE 3.80e-01
 US-10-029-386-31530

Query Match 74.6%; Score 47; DB 14; Length 181;
 Best Local Similarity 66.7%; Pred. No. 77;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWWPTCL 9
 |||||
 Db 102 SPRWWVACM 110

RESULT 5
 US-10-029-386-31685
 ; Sequence 31685, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: AEOMICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 31685
 ; LENGTH: 34
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC010422.6
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
 US-10-029-386-31685

Query Match 73.0%; Score 46; DB 14; Length 34;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWNP 7
 |||||
 Db 14 PRWNP 19

RESULT 6
 US-09-764-869-648
 ; Sequence 648, Application US/09764869
 ; Patent No. US20020061521A1
 ; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 648
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-869-648

Query Match 73.0%; Score 46; DB 9; Length 72;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPTCL 9
| | | | |
Db 49 RWPPTCL 55

RESULT 7
US-10-091-504-648
; Sequence 648, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 648
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-504-648

Query Match 73.0%; Score 46; DB 14; Length 72;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPTCL 9
| | | | |
Db 49 RWPPTCL 55

RESULT 8
US-10-227-577-648
; Sequence 648, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 648
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-227-577-648

Query Match 73.0%; Score 46; DB 15; Length 72;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPPTCL 9
| | | | |
Db 49 RWPPTCL 55

RESULT 9
US-10-104-047-3255
; Sequence 3255, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3255
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3255

Query Match 73.0%; Score 46; DB 15; Length 256;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SPRWPTC B
| | | | |
Db 211 SSRWPAC 218

RESULT 10
US-10-437-963-152726
; Sequence 152726, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152726
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(371)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52749C.1.pgp
US-10-437-963-152726

Query Match          71.4%; Score 45; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPRWMP 6
Db 79 SPRWMP 84

RESULT 11
US-10-437-963-159523
; Sequence 159523, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159523
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58891C.1.pgp
US-10-437-963-159523

Query Match          69.8%; Score 44; DB 16; Length 102;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWWPTCL 9
Db 37 RWWATCL 43

RESULT 12
US-10-394-575-81
; Sequence 81, Application US/10394575
; Publication No. US20030236393A1
; GENERAL INFORMATION:
; APPLICANT: TRUCKSIS, MICHELE
; TITLE OF INVENTION: VIRULENCE GENES OF M. MARINUM AND M. TUBERCULOSIS
; FILE REFERENCE: VET-2
; CURRENT APPLICATION NUMBER: US/10/394,575
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/367,206
```

```
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/366,262
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Mycobacterium marinum
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (57)
; OTHER INFORMATION: variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (197)
; OTHER INFORMATION: variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (206)
; OTHER INFORMATION: variable amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (212)
; OTHER INFORMATION: variable amino acid
US-10-394-575-81

Query Match          69.8%; Score 44; DB 15; Length 248;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPRWPTC 8
Db 26 SRRWWPPC 33

RESULT 13
US-09-733-183A-20
; Sequence 20, Application US/09733183A
; Patent No. US20020081568A1
; GENERAL INFORMATION:
; APPLICANT: DelVecchio, Alfred M.
; TITLE OF INVENTION: Hepatitis C Virus NS5B Truncated Protein
; FILE REFERENCE: P50743D1
; CURRENT APPLICATION NUMBER: US/09/733,183A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/208140
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069208
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-09-733-183A-20

Query Match          68.3%; Score 43; DB 9; Length 49;
Best Local Similarity 75.0%; Pred. No. 97;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWPTCL 9
Db 27 PRWFFCL 34

RESULT 14
US-10-342-372-20
; Sequence 20, Application US/10342372
; Publication No. US20030190606A1
; GENERAL INFORMATION:
```

; APPLICANT: DelVecchio, Alfred M.
; TITLE OF INVENTION: Hepatitis C Virus NS5B Truncated Protein
; FILE REFERENCE: P50743D1
; CURRENT APPLICATION NUMBER: US/10/342,372
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/733,183A
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/208140
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069208
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Viral
US-10-342-372-20

Query Match 68.3%; Score 43; DB 14; Length 49;
Best Local Similarity 75.0%; Pred. No. 97;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPTCL 9
Db 27 PRWPTCL 34
|||.|||

RESULT 15
US-10-112-944-429
; Sequence 429, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: Pt_Fl_Genes Version 5.0
; SEQ ID NO 429
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-112-944-429

Query Match 68.3%; Score 43; DB 12; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWPTC 8
Db 99 WWPTC 103
|||||

Search completed: September 13, 2004, 14:38:50
Job time : 11.2105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:10:50 ; Search time 2.52632 Seconds
(without alignments)
342.682 Million cell updates/sec

Title: US-10-019-219A-2

Perfect score: 63

Sequence: 1 SPRWPTCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	806	2 G84863	hypothetical prote
2	43	68.3	120	2 T42054	hypothetical prote
3	42	66.7	221	2 C81099	hypothetical prote
4	42	66.7	221	2 C81842	hypothetical prote
5	42	66.7	275	2 I40211	probable sterol de
6	42	66.7	443	2 I39538	alpha-amylase - Ae
7	42	66.7	2092	2 S30026	genome polyprotein
8	42	66.7	2149	2 S18676	genome polyprotein
9	41	65.1	123	2 H72698	hypothetical prote
10	41	65.1	252	1 JQ0417	oleoyl-[acyl-carri
11	41	65.1	295	2 AD3577	sugar transport sy
12	41	65.1	461	2 H70899	probable cytochrom
13	41	65.1	790	2 T34293	hypothetical prote
14	41	65.1	1398	2 T18350	probable pol polyp
15	40	63.5	64	2 S28486	hypothetical prote
16	40	63.5	212	2 S74376	hypothetical prote
17	40	63.5	263	1 S43189	hypothetical prote
18	40	63.5	290	1 D47468	cytochrome-c oxida
19	40	63.5	471	2 T50016	transcription fact
20	40	63.5	542	2 I39540	chitinase (EC 3.2.
21	40	63.5	997	2 T39521	hypothetical signa
22	39	61.9	119	2 A53257	H+-transporting AT
23	39	61.9	176	2 H75332	ankyrin-related pr
24	39	61.9	338	2 A87566	sensor histidine k
25	39	61.9	376	2 S33654	zinc transport pro
26	39	61.9	567	2 E96764	ethylene-insensiti
27	38.5	61.1	423	2 AG2394	hypothetical prote
28	38	60.3	170	2 S26718	hypothetical 19.8K
29	38	60.3	175	2 T27543	hypothetical prote

30	38	60.3	197	2 S59397	probable membrane
31	38	60.3	211	1 MKAD	lysozyme (EC 3.2.1
32	38	60.3	248	2 S23449	NADH oxidase (H2O2
33	38	60.3	335	2 T46351	hypothetical prote
34	38	60.3	369	2 T40279	hypothetical prote
35	38	60.3	374	2 T43708	cytochrome oxidase
36	38	60.3	379	2 E83597	homoserine O-acety
37	38	60.3	528	2 E75310	conserved hypothet
38	38	60.3	584	2 B84668	ethylene-insensiti
39	38	60.3	598	2 C82194	ATP-dependent prot
40	38	60.3	731	2 JC7701	ARGAP9 protein -
41	38	60.3	770	2 C87316	alpha-N-acetylgluc
42	38	60.3	1182	2 I48378	hairless protein -
43	37	58.7	133	2 F72471	hypothetical prote
44	37	58.7	149	2 A87346	hypothetical prote
45	37	58.7	247	2 A70910	probable truncated

ALIGNMENTS

RESULT 1

G84863

hypothetical protein At2g43240 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: G84863

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84863

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-806 <STO>

A/Cross-references: GB:AE002093; MID:g3763933; PIDN:AA64313.1; GSPDB:GN00139

C/Genetics:

A/Gene: At2g43240

A/Map position: 2

Query Match 73.0%; Score 46; DB 2; Length 806;

Best Local Similarity 75.0%; Pred. No. 17;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 2 PRWPTCL 9

|:|||||

Db 676 PKWPTSL 683

RESULT 2

T42054

hypothetical protein - Streptomyces coelicolor (fragment)

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T42054

R/Guajarro, J.; Santamaria, R.; Schauer, A.; Losick, R.

J. Bacteriol. 170, 1895-1901, 1988

A/Title: Promoter determining the timing and spatial localization of transcription of a c

A/Reference number: 222043; MUID:88169521; PMID:2450872

A/Accession: T42054

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-120 <GUI>

A/Cross-references: EMBL:M20145; PIDN:AA26812.1

Query Match 68.3%; Score 43; DB 2; Length 120;

Best Local Similarity 100.0%; Pred. No. 8.6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 4 WWPTC 8

|||||

```

Db      88  RWPTC 92

RESULT 3
C81099
hypothetical protein NMB1296 [imported] - Neisseria meningitidis (strain MCS8 serogroup
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: C81099
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: C81099
A:Status: preliminary
A:Gene: NMB1296
A:Molecule type: DNA
A:Residues: 1-221 <TET>
A:Cross-references: GB:AE002478; GB:AE002098; NID:G7226533; PIDN:AAF41672.1; PID:G722653
A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1296

Query Match      66.7%; Score 42; DB 2; Length 221;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  RWPTC 8
      |||||
Db      177  RWPTC 182

RESULT 4
C81842
hypothetical protein NMA1506 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: C81842
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: C81842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; NID:G7380091; PIDN:CAB84738.1; PID:G738015
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1506

Query Match      66.7%; Score 42; DB 2; Length 221;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  RWPTC 8
      |||||
Db      177  RWPTC 182

RESULT 5
I40211
probable sterol dehydrogenase (EC 1.1.1.-) - Bradyrhizobium japonicum
C:Species: Bradyrhizobium japonicum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999
C:Accession: I40211
R:Tully, R.E.; Keister, D.L.
Appl. Environ. Microbiol. 59, 4136-4142, 1993
A:Title: Cloning and mutagenesis of a cytochrome P-450 locus from Bradyrhizobium japonic
A:Reference number: I40207
A:Accession: I40211
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-275 <RES>
A:Cross-references: EMBL:U12678; NID:G529961; PIDN:AAC28892.1; PID:G529965
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: oxidoreductase
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match      66.7%; Score 42; DB 2; Length 275;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2  PRWPTC 8
      |||||
Db      219  PRWPTC 225

RESULT 6
I39538
alpha-amylase - Aeromonas hydrophila
C:Species: Aeromonas hydrophila
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 03-Dec-1999
C:Accession: I39538
R:Chang, M.C.; Chang, J.C.; Chen, J.P.
J. Gen. Microbiol. 139, 3215-3223, 1993
A:Title: Cloning and nucleotide sequence of an extracellular alpha-amylase gene from Aer
A:Reference number: I39538; MUID:94172314; PMID:8126440
A:Accession: I39538
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-443 <RES>
A:Cross-references: GB:LI9299; NID:G304014; PIDN:AAA21016.1; PID:G304015
C:Genetics:
A:Gene: amyA
C:Superfamily: mammalian alpha-amylase; alpha-amylase core homology
F:166-291/Domain: alpha-amylase core homology <AMY>

Query Match      66.7%; Score 42; DB 2; Length 443;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  RWPTC 8
      |||||
Db      406  RWPTC 411

RESULT 7
S30026
genome polypeptide - Rift Valley fever virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: Rift Valley fever virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S30026
R:Muller, R.; Argentin, C.; Bouloy, M.; Prehaud, C.; Bishop, D.H.L.
Nucleic Acids Res. 20, 6440, 1992
A:Title: Corrigendum: Completion of the genome sequence of Rift Valley fever phlebovirus
A:Reference number: S30026
A:Accession: S30026
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-2092 <MUL>
A:Cross-references: GB:X56464
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
C:Superfamily: Uukuniemi virus RNA-directed RNA polymerase
C:Keywords: nucleotidyltransferase

Query Match      66.7%; Score 42; DB 2; Length 2092;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  SPRWPTC 6
      |||||

```

Db 1013 SPKWWP 1018

RESULT 8
S18676
genome polyprotein - Rift Valley fever virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Rift Valley fever virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S18676
R;Muller, R.; Argentin, C.; Bouloy, M.; Prehaud, C.; Bishop, D.H.L.
Nucleic Acids Res. 19, 5433, 1991
A;Title: Completion of the genome sequence of Rift Valley fever phlebovirus indicates the
A;Reference number: S18676; MUID: 92020238; PMID: 1923828
C;Accession: S18676
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-2149 <MUL>
A;Cross-references: EMBL:X56464; NID:g61926; PIDN:CAA39836.1; PID:g61927
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: Ukuniemi virus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase

Query Match 66.7%; Score 42; DB 2; Length 2149;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPKWWP 6

Db 1013 SPKWWP 1018

RESULT 9
H72698
hypothetical protein APE1008 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C;Accession: H72698
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID: 99310339; PMID: 10382966
C;Accession: H72698
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-123 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79992.1; PID:gi043778; PID:g510
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1008
C;Superfamily: Aeropyrum pernix hypothetical protein APE1008

Query Match 65.1%; Score 41; DB 2; Length 123;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWWPTCL 9

Db 110 RWWPTCL 116

RESULT 10
JQ0417
oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) homolog - Vibrio anguillarum
N;Alternate names: hypothetical 28.1K protein; ORF6 protein; S-acyl fatty acid synthase
C;Species: Vibrio anguillarum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2003
C;Accession: JQ0417; S26422
R;Farrell, D.H.; Mikesell, P.; Actis, L.A.; Crosa, J.H.
Gene 86, 45-51, 1990
A;Title: A regulatory gene, angR, of the iron uptake system of Vibrio anguillarum: simil
A;Reference number: JQ0416; MUID: 90185247; PMID: 2311935

A;Accession: JQ0417
A;Molecule type: DNA
A;Residues: 1-252 <FAR>
A;Cross-references: GB:M34504; NID:gi55150; PIDN:AAA79861.1; PID:gi55153
R;Tolmasky, M.E.; Actis, L.A.; Waldbeser, L.S.; Crosa, J.H.
submitted to the EMBL Data Library, April 1992
A;Description: Genetic characterization of the regulatory protein AngR: presence of leuc
A;Reference number: S26421
A;Accession: S26422
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-252 <TOL>
A;Cross-references: EMBL:Z12000; NID:g48322; PIDN:CAA78045.1; PID:g48324
C;Comment: This pathogenic bacterium is a causative agent of vibriosis, a widespread sept
C;Superfamily: thioesterase, type II; oleoyl-[acyl-carrier-protein] hydrolase homology
C;Keywords: thioester hydrolase
F;19-234/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 65.1%; Score 41; DB 1; Length 252;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRWPTCL 9

Db 156 SPEWWPIFL 164

RESULT 11

AD3577

sugar transport system permease protein [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AD3577
R;DeIvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensi
A;Reference number: AD35252; PMID: 11756688
A;Accession: AD3577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53783.1; PID:gi17984712; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10541
A;Map position: 11
C;Superfamily: maltose transport protein malG

Query Match 65.1%; Score 41; DB 2; Length 295;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWPTCL 7

Db 63 SPKWWPS 69

RESULT 12

H70899

probable cytochrome P450 Rv1394c [similarity] - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C;Accession: H70899
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome s
A;Reference number: A70500; MUID: 98295987; PMID: 9634230
A;Accession: H70899
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-461 <COL>
 A:Cross-references: GB:Z80108; GB:AL123456; NID:G3256012; PIDN:CAB02176.1; PID:G1542902
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV1394C
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
 C:Keywords: heme; iron; metalloprotein
 F:271-431/Domain: cytochrome P450 homology <P45>
 F:409/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 65.1%; Score 41; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWNP 6
 DB 373 PRWNP 377

RESULT 13
 T34293
 Hypothetical protein F49E10.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T34293
 R:Miller, N.
 submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid F49E10.
 A:Reference number: Z21500
 A:Accession: T34293
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-790 <MIL>
 A:Cross-references: EMBL:U53341; PIDN:AACG9106.1; GSPDB:GN000028; CESP:F49E10.2
 A:Experimental source: strain Bristol N2; clone F49E10
 C:Genetics:
 A:Gene: CESP:F49E10.2
 A:Map position: X
 A:Introns: 51/1; 92/3; 176/3; 235/3; 332/2; 514/1; 543/2; 569/3; 677/1; 732/3

Query Match 65.1%; Score 41; DB 2; Length 790;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWNP 6
 DB 688 PRWNP 692

RESULT 14
 T18350
 Probable polypeptide - rice blast fungus gypsy retroelement (fragment)
 C:Species: Magnaporthe grisea (rice blast fungus)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18350
 R:Dobinson, K.F.
 submitted to the EMBL Data Library, September 1994
 A:Description: Sequence of the gfp retroelement.
 A:Reference number: Z18883
 A:Accession: T18350
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1398 <DOB>
 A:Cross-references: EMBL:M77661; NID:G538065; PID:G538067; PIDN:AAA21442.1
 C:Genetics:
 A:Mobile element: gypsy retroelement

Query Match 65.1%; Score 41; DB 2; Length 1398;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWNP 6

DB 76 PRWNP 80

RESULT 15

S28486
 Hypothetical protein 2 - Vibrio cholerae
 C:Species: Vibrio cholerae
 C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
 C:Accession: S28486
 R:Manning, P.A.
 submitted to the EMBL Data Library, May 1991
 A:Reference number: S28467
 A:Accession: S28486
 A:Molecule type: DNA
 A:Residues: 1-64 <MAN>
 A:Cross-references: EMBL:X59554; NID:G48381; PIDN:CAA42152.1; PID:G48401
 A:Experimental source: strain 017

Query Match 63.5%; Score 40; DB 2; Length 64;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWNP 8

DB 51 SYRWNP 58

Search completed: September 13, 2004, 14:23:16
 Job time : 4.52632 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 13:48:45 ; Search time 1.52632 Seconds
(without alignments)
307.034 Million cell updates/sec

Title: US-10-019-219A-2
Perfect score: 63
Sequence: 1 SPRWWPTCL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	68.3	817	1 DLG3 HUMAN	Q92796 homo sapien
2	43	68.3	849	1 DLG3 RAT	Q62936 rattus norv
3	42	66.7	443	1 AMYA_AERHY	P41131 aeromonas h
4	42	66.7	849	1 DLG3 MOUSE	P70175 mus musculu
5	42	66.7	2149	1 RRPL_RPVFVZ	P27316 rift valley
6	41	65.1	252	1 SAST_VIBAN	P19829 vibrio angu
7	41	65.1	461	1 C132 MYCBO	P59954 mycobacteri
8	41	65.1	461	1 C132 MYCTU	P77900 mycobacteri
9	40	63.5	471	1 EIL4 ARATH	Q91x16 arabidopsis
10	40	63.5	518	1 EIL2 ARATH	Q23115 arabidopsis
11	39	61.9	119	1 ATP6_NARPO	P22067 naegleria f
12	39	61.9	376	1 ZRT1_YEAST	P32804 saccharomyc
13	39	61.9	567	1 EIL3 ARATH	Q23116 arabidopsis
14	38	60.3	175	1 YOCA CAEEL	Q23280 caenorhabdi
15	38	60.3	211	1 LYCH CHASP	P00721 chalaropsis
16	38	60.3	374	1 OX11_SCHPO	Q14300 schizosacch
17	38	60.3	379	1 METX_PSEAE	P57714 pseudomonas
18	38	60.3	417	1 O85E_DROME	P81924 drosophila
19	38	60.3	584	1 EIL1 ARATH	Q9slh0 arabidopsis
20	38	60.3	896	1 LX23 HORVU	Q8gsm2 tartus norv
21	38	60.3	1181	1 HAIR RAT	P97609 rattus norv
22	38	60.3	1182	1 HAIR MOUSE	Q61645 mus musculu
23	38	60.3	1189	1 HAIR HUMAN	Q43593 homo sapien
24	37	58.7	309	1 ER25_YEAST	P33045 saccharomyc
25	37	58.7	311	1 LUCI RENRE	P27652 renilla ren
26	37	58.7	418	1 VIAR SHEEP	P48043 ovine aries
27	37	58.7	460	1 YS85 MYCTU	Q10809 mycobacteri
28	37	58.7	546	1 LNT_TREPA	Q83279 treponema p
29	37	58.7	809	1 YATA_SCHPO	Q10155 schizosacch
30	37	58.7	1131	1 YANC SCHPO	Q10077 schizosacch
31	37	58.7	1385	1 YMS5 CAEEL	P34501 caenorhabdi
32	36.5	57.9	226	1 MDCG BRAJA	Q89xp2 bradyrhizob
33	36.5	57.9	830	1 HMT1_SCHPO	Q02532 schizosacch

34 36 57.1 94 1 YOR6 MNV
35 36 57.1 238 1 RS4E PYRAE
36 36 57.1 283 1 ISPE CHLMU
37 36 57.1 333 1 A85C MYCLE
38 36 57.1 340 1 A85C MYCTU
39 36 57.1 352 1 A85C MYCAV
40 36 57.1 431 1 CIT1_ECOLI
41 36 57.1 434 1 CIT1_SALTY
42 36 57.1 696 1 YIK8 YEAST
43 36 57.1 790 1 ATSY_SYNPF
44 36 57.1 1026 1 BGAL_STRTR
45 36 57.1 1165 1 POL_GALV

P15099 narcissus m
Q8ztd3 pyrobaculum
Q9plc0 chlamydia m
Q05862 mycobacteri
P31953 mycobacteri
O52972 mycobacteri
P07661 escherichia
P24115 salmonella
P40483 saccharomyc
P37385 synechococc
P23989 streptococc
P21414 gibbon ape

ALIGNMENTS

RESULT 1
DLG3_HUMAN STANDARD; PRT; 817 AA.
AC Q92796; Q9ULI8;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Presynaptic protein SAP102 (Synapse-associated protein 102)
DE (Neuroendocrine-DLG) (NE-DLG) (Discs, large homolog 3).
GN DLG3 OR KIAA1232.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97332623; PubMed=9188857;
RA Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T.,
RA Sasaki J., Nakao M., Kuwano A., Nakata M., Ushio Y., Saya H.;
RT "Cloning and characterization of NE-dlg: a novel human homolog of the
RT Drosophila discs large (dlg) tumor suppressor protein interacts with
RT the APC protein.";
RL Oncogene 14:2425-2433 (1997).
RN [2]
RP SEQUENCE OF 330-817 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345 (1999).
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
CC SUBUNIT NR2B (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the MAGUK family.
CC -!- SIMILARITY: Contains 3 PDZ/DHR domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U49089; AAB61453.1; -;
DR EMBL; AB033058; BAA86546.1; -;
DR HSSP; Q12959; 1PDR.
DR Genew; HGNC:2902; DLG3.
DR TM; 300189; -;
DR GO; GO:0004385; F:guanylate kinase activity; NAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; NAS.

```

DR InterPro; IPR008144; Guanylate_kin.
DR InterPro; IPR008145; Guanylt/Ca.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00072; GuKC; 1.
DR SMART; SM00228; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS0052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain; Repeat.
KW SH3 domain; Repeat.
FT DOMAIN 130 217 PDZ 1.
FT DOMAIN 226 311 PDZ 2.
FT DOMAIN 379 465 PDZ 3.
FT DOMAIN 503 568 SH3.
FT DOMAIN 628 803 GUANYLATE_KINASE.
FT CONFLICT 330 381
FT FT
FT FT
FT CONFLICT 592 606
FT FT
FT SEQUENCE 817 AA; 90344 MW; 3D7512EC4713FC4E CRC64;

Query Match 68.3%; Score 43; DB 1; Length 817;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRWWPTC 8
Db 103 PSWWPEC 109

RESULT 2
DLG3_RAT
ID DLG3_RAT STANDARD; PRT; 849 AA.
AC Q62936; P70547;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Presynaptic protein SAP102 (synapse-associated protein 102) (PSD-
95/SAP90 related protein 1) (Discs, large homolog 3).
GN DLG3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=96374358; PubMed=8780649;
RA Mueller B.M., Kistner U., Kindler S., Chung W.J., Kuhlendahl S.,
RA Fenster S.D., Lau L.-F., Veh R.W., Haganir R.L., Gundelfinger E.D.,
RA Garner C.C.;
RT "SAP102, a novel postsynaptic protein that interacts with NMDA
RT receptor complexes in vivo.";
RL Neuron 17:255-265(1996).
RN [2]
SEQUENCE FROM N.A. (ISOFORM SHORT).
RA Irie M., Hata Y., Takai Y.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
CC SUBUNIT NR2B.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q62936-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q62936-2; Sequence=VSP_003151;
CC

CC -!- SIMILARITY: Belongs to the MAGUK family.
CC -!- SIMILARITY: Contains 3 PDZ/DHR domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U50147; AAA93031.1; -.
CC EMBL; U53367; AAB48561.1; -.
CC HSP; Q12959; IPDR.
DR InterPro; IPR008144; Guanylate_kin.
DR InterPro; IPR008145; Guanylt/Ca.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00072; GuKC; 1.
DR SMART; SM00228; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS0052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain; Repeat; Alternative splicing.
KW SH3 domain; Repeat; Alternative splicing.
FT DOMAIN 149 235 PDZ 1.
FT DOMAIN 244 330 PDZ 2.
FT DOMAIN 404 484 PDZ 3.
FT DOMAIN 519 589 SH3.
FT DOMAIN 659 849 GUANYLATE_KINASE.
FT VARSPLIC 627 640 Missing (in isoform Short).
FT /FTID=VSP_003151.
SQ SEQUENCE 849 AA; 93539 MW; 34DA9C46C7BB96DB CRC64;

Query Match 68.3%; Score 43; DB 1; Length 849;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PRWWPTC 8
Db 103 PSWWPEC 109

RESULT 3
AMYA_AERHY
ID AMYA_AERHY STANDARD; PRT; 443 AA.
AC P41131;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
OC Aeromonadaceae; Aeromonas.
OX NCBI_TaxID=644;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MCC-1;
RX MEDLINE=94172314; PubMed=8126440;
RA Chang M.C., Chang J.C., Chen J.P.;
RT "Cloning and nucleotide sequence of an extracellular alpha-amylase
RT gene from Aeromonas hydrophila MCC-1";
RL J. Gen. Microbiol. 139:3215-3223(1993).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic

```

```
CC linkages in oligosaccharides and polysaccharides.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L19299; AAA21016.1; -.
CC PIR; I39538; I39538.
CC HSSP; P29957; 1AQM.
CC InterPro; IPR006589; Alp_aml1_cat_sub.
CC InterPro; IPR006048; Alpha_aml1_C.
CC InterPro; IPR006047; Alpha_aml1_cat.
CC InterPro; IPR006046; Glyco_hydro_13.
CC Pfam; PF00128; alpha-amy1ase; 1.
CC Pfam; PF02806; alpha-amy1ase; C; 1.
CC PRINTS; PR00110; ALPHAAMYLASE.
CC SMART; SM00642; Amy; 1.
CC Hydrolase; Glycosidase; Carbohydrate metabolism; Signal.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 443 ALPHA-AMYLASE.
CC ACT_SITE 198 198 BY SIMILARITY.
CC ACT_SITE 202 202 BY SIMILARITY.
CC ACT_SITE 287 287 BY SIMILARITY.
CC SEQUENCE 443 AA; 48333 MW; 8F8D60B9341A92F9 CRC64;
CC
CC Query Match 66.7%; Score 42; DB 1; Length 443;
CC Best Local Similarity 83.3%; Pred. No. 19;
CC Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 3 RWPPTC 8
CC |||||
CC Db 406 RWPPTC 411
CC
CC RESULT 4
CC DLG3 MOUSE
CC ID DLG3 MOUSE STANDARD; PRT; 849 AA.
CC AC P70175;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Presynaptic protein SAP102 (Synapse-associated protein 102) (Discs,
CC large homolog 3).
CC GN DLG3 OR DLGH3.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC [1]
CC SEQUENCE FROM N.A.
CC SPRAIN-CS7BL/6; TISSUE=Brain;
CC RA Kohmura N., Makino S., Yagi T.;
CC RA Submittion (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
CC SUBUNIT NR2B.
CC
CC -!- SIMILARITY: Belongs to the MAGUK family.
CC -!- SIMILARITY: Contains 3 PDZ/DHR domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```
CC EMBL; D87117; BAA13249.1; -.
CC HSSP; Q12959; 1PDR.
CC MGD; MGI:1888986; Dlg3.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0005515; F:protein binding; IPI.
CC InterPro; IPR008144; Guanylate_kin.
CC InterPro; IPR008145; Guanylt/Ca.
CC InterPro; IPR001478; PDZ.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00625; Guanylate_kin; 1.
CC Pfam; PF00595; PDZ; 3.
CC Pfam; PF00018; SH3; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00072; GuKC; 1.
CC SMART; SM00228; PDZ; 3.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE; PS00502; GUANYLATE_KINASE_2; 1.
CC PROSITE; PS0106; PDZ; 3.
CC PROSITE; PS00002; SH3; 1.
CC SH3 domain; Repeat.
CC FT DOMAIN 149 235 PDZ 1.
CC FT DOMAIN 244 330 PDZ 2.
CC FT DOMAIN 404 484 PDZ 3.
CC FT DOMAIN 519 589 SH3.
CC FT DOMAIN 659 849 GUANYLATE_KINASE.
CC SQ SEQUENCE 849 AA; 93482 MW; EF3EF2D7513538EE CRC64;
CC
CC Query Match 66.7%; Score 42; DB 1; Length 849;
CC Best Local Similarity 71.4%; Pred. No. 35;
CC Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 2 PRWPTC 8
CC |||||
CC Db 103 PGWPTC 109
CC
CC RESULT 5
CC RRPL RVFVZ
CC ID RRPL RVFVZ STANDARD; PRT; 2149 AA.
CC AC P27316;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
CC GN L.
CC OS Rift valley fever virus (strain ZH-548 M12) (RVFV).
CC OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
CC NCBI_TaxID=11589;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=92020238; PubMed=1923828;
CC Mueller R., Argentin C., Bouloy M., Prehaud C., Bishop D.H.L.;
CC RA "Completion of the genome sequence of Rift Valley fever phlebovirus
CC indicates that the L RNA is negative sense and codes for a putative
CC transcriptase-replicase."
CC RT Nucleic Acids Res. 19:5433-5433(1991).
CC RL -!- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC {RNA} (N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56464; CAA39836.1; -.
CC PIR; S18676; S18676.
CC InterPro; IPR007322; Bunya_RdRp.
CC InterPro; IPR007099; RNA_pol_NSvir.
CC
```

DR Pfan; PF04196; Bunya RdRp; 1.
 KW Transferase; Nucleotidyltransferase; RNA-directed RNA polymerase.
 SQ SEQUENCE 2149 AA; 243589 MW; 8D5739C6079A88D7 CRC64;

Query Match 66.7%; Score 42; DB 1; Length 2149;
 Best Local Similarity 83.3%; Pred. No. 86;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPRWMP 6
 |||
 Db 1013 SPKWMP 1018

RESULT 6
 SAST_VIBAN STANDARD; PRT; 252 AA.
 AC P19629;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable anguibactin biosynthesis thioesterase (EC 3.1.2.-).
 OS Vibrio anguillarum (Listonella anguillarum).
 OG plasmid pJMI.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Listonella.
 OX NCBI_TaxID=55601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=775;
 RX MEDLINE=90185247; PubMed=2311935;
 RA Farrell D.H., Mikesell P., Actis L.A., Crosa J.H.;
 RT "A regulatory gene, angr, of the iron uptake system of Vibrio
 anguillarum: similarity with phage P22 cro and regulation by iron";
 RL Gene 86:45-51(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=531A;
 RX MEDLINE=93328275; PubMed=8335354;
 RA Tolmasey M.E., Actis L.A., Crosa J.H.;
 RT "A single amino acid change in Angr, a protein encoded by pJMI-like
 RT virulence plasmids, results in hyperproduction of anguibactin";
 RL Infect. Immun. 61:3228-3233(1993).
 CC -!- FUNCTION: Probable thioesterase.
 CC -!- PATHWAY: Anguibactin siderophore biosynthesis.
 CC -!- SIMILARITY: TO OTHER THIOESTERASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M34504; AAA79861.1; --
 DR EMBL; Z12000; CAA78045.1; --
 DR FIR; JQ0417; JQ0417.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00975; Thioesterase; 1.
 KW Plasmid; Hydrolase.
 FT ACT SITE 92 92 BY SIMILARITY.
 FT ACT SITE 229 229 BY SIMILARITY.
 FT ACT SITE 252 252 BY SIMILARITY.
 SQ SEQUENCE 252 AA; 28070 MW; 1FB1AA3CCEDB99F4 CRC64;

Query Match 65.1%; Score 41; DB 1; Length 252;
 Best Local Similarity 86.7%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPRWMP 9
 |||
 Db 156 SPEWMP 164

RESULT 7
 CL32_MYCBO STANDARD; PRT; 461 AA.
 AC P59954;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Putative cytochrome P450 132 (EC 1.14.-.-).
 GN CYP132 OR ME1429C.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972;
 RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Farkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; BX248338; CAD94290.1; --
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
 KW Complete proteome.
 FT METAL 409 409 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 461 AA; 52186 MW; EA17E6EAEAA05791 CRC64;

Query Match 65.1%; Score 41; DB 1; Length 461;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PRWMP 6
 |||
 Db 373 PRWMP 377

RESULT 8
 CL32_MYCTU STANDARD; PRT; 461 AA.
 AC P77900;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative cytochrome P450 132 (EC 1.14.-.-).
 GN CYP132 OR RV1394C OR MT1439 OR MTCY21B4.11C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R.,

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RL complete genome sequence.";
 RN Nature 393:537-544(1998).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 EX MEDLINE=22206494; PubMed=12218036;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL J. Bacteriol. 184:5479-5490(2002).
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z80108; CAB02176.1; --.
 CC EMBL; AE007015; AAK45704.1; --.
 CC PIR; H70899; H70899.
 CC HSSP; P14779; 1JPZ.
 CC TIGR; MT1439; --.
 CC Tuberculist; Rv1394c; --.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC PRINTS; PF00067; P450.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
 CC Complete proteome.
 CC METAL 409 409 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 135 135 R -> L (IN REF. 2).
 FT SEQUENCE 461 AA; 52229 MW; 2DEF61C8A10B0CF3 CRC64;
 Query Match 65.1%; Score 41; DB 1; Length 461;
 Best Local Similarity 100.0%; Pred.No. 28;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 PRWVP 6
 Db 373 PRWVP 377
 RESULT 9
 ID EIL4 ARATH STANDARD; PRT; 471 AA.
 AC Q9LX16;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Putative ETHYLENE-INSENSITIVE3-like 4 protein.
 GN EIL4 OR AT5G10120 OR T31P16.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 OC [1]
 RN SEQUENCE FROM N.A.
 RP STEAIN=cv. Columbia;
 RC MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsumoto A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,

RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Fohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
 RA Feldzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Weidpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rued S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:823-826(2000).
 CC -!- FUNCTION: Putative transcription factor that may be involved in
 CC the ethylene response pathway (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the EIN3 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL356332; CAB92053.1; --.
 CC PIR; T50016; T50016.
 CC InterPro; IPR006957; EIN3.
 CC Pfam; PF04873; EIN3; 1.
 CC KW Hypothetical protein; Transcription regulation; Nuclear protein.
 CC SEQUENCE 471 AA; 53954 MW; 864B80722F6B5A32 CRC64;
 Query Match 63.5%; Score 40; DB 1; Length 471;
 Best Local Similarity 71.4%; Pred.No. 40;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 SPRWVPT 7
 Db 188 APPWVPT 194
 RESULT 10
 ID EIL2 ARATH STANDARD; PRT; 518 AA.
 AC O23115;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ETHYLENE-INSENSITIVE3-like 2 protein.
 GN EIL2 OR AT5G1120 OR T10P18.150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 OC [1]
 RN SEQUENCE FROM N.A., AND FUNCTION.
 RP STRAIN=cv. Columbia;
 RC MEDLINE=97358539; PubMed=9215635;
 RA Chao O., Rothenberg M., Solano R., Roman G., Terraghi W., Ecker J.R.;
 RA "Activation of the ethylene gas response pathway in Arabidopsis by the
 RT nuclear protein ETHYLENE-INSENSITIVE3 and related proteins.";

```

RL Cell 89:1133-1144 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naro K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spith J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney J., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA Patrell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchhoff K., Roth K., King L., Bahret A., Miller B., Marra M.A.,
RA Martensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkee W., Mooijman P., Klein Lankhorst R.,
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Berniseier S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:823-826 (2000).
RN [3]
RP CHARACTERIZATION, AND FUNCTION.
RX MEDLINE=99069218; PubMed=9851977;
RA Solano R., Stepanova A.N., Chao Q., Ecker J.R.;
RT "Nuclear events in ethylene signaling: a transcriptional cascade
RT mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1.";
RL Genes Dev. 12:3703-3714 (1998).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=22506420; PubMed=12606727;
RA Alonso J.M., Stepanova A.N., Solano R., Wisman E., Ferrari S.,
RA Ausubel F.M., Ecker J.R.;
RT "Five components of the ethylene-response pathway identified in a
RT screen for weak ethylene-insensitive mutants in Arabidopsis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:2992-2997 (2003).
CC -!- FUNCTION: Probable transcription factor acting as a positive
CC regulator in the ethylene response pathway. Could bind the primary
CC ethylene response element present in the ETHYLENE-RESPONSE-FACTOR1
CC promoter.
CC -!- SUBUNIT: Acts as homodimer to bind the primary ethylene response
CC element (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the EIN3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF004214; AAC49747.1; -.
CC DR EMBL; AC140977; AAC073887.1; -.
CC DR TRANSFAC; T02651; -.
CC DR InterPro; IPR006957; EIN3.
CC DR Pfam; PF04873; EIN3; 1.
CC KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC DOMAIN
CC Coiled coil.
CC SEQUENCE 518 AA; 59185 MW; 173EA49BE9A17689 CRC64;
CC -----
Query Match 63.5%; Score 40; DB 1; Length 518;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SPRWNP 7
DB 214 TTPWNP 220
RESULT 11
ATP6_NAEFO STANDARD; PRT; 119 AA.
ID ATP6_NAEFO
AC P22067;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase A chain (EC 3.6.3.14) (Protein 6) (Fragment).
GN ATP6 OR OLI2.
OS Naegleria fowleri.
OG Mitochondrion.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCBI_TaxID=5763;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LEE;
RC MEDLINE=91178040; PubMed=2007628;
RA McLaughlin G.L., Vodkin M.H., Huizinga H.W.;
RT "Amplification of repetitive DNA for the specific detection of
RT Naegleria fowleri.";
RL J. Clin. Microbiol. 29:227-230 (1991).
CC -!- FUNCTION: Key component of the proton channel; it may play a
CC direct role in the translocation of protons across the membrane.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ATPase A chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55009; CAB25936.1; -.
CC DR PIR; A53257; A53257.
CC DR InterPro; IPR000568; ATPsynth_Asub.
CC DR Pfam; PF00119; ATP-synt_A; 1.
CC DR PRINTS; PR00123; ATPASEA.
CC DR PROSITE; PS00449; ATPASE_A; PARTIAL.
CC KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
CC FT NON_TER 1
CC FT NON_TER 119
CC SQ SEQUENCE 119 AA; 13934 MW; 50892FC0BB5C04F5 CRC64;
CC -----
Query Match 61.9%; Score 39; DB 1; Length 119;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 WNPTEL 9
DB 70 WNPTEL 75
RESULT 12
ZRT1_YEAST STANDARD; PRT; 376 AA.
ID ZRT1_YEAST

```


Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 Arakawa T., Banh J., Banno F., Bowser L., Brooks S.V., Carninci P.,
 Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.:
 "Empirical analysis of transcriptional activity in the Arabidopsis
 genome".
 Science 302:842-846 (2003).
 [4]
 RP CHARACTERIZATION, AND FUNCTION.
 RX MEDLINE=99069218; PubMed=9851977;
 RY Solano R., Stepanova A.N., Chao Q., Ecker J.R.:
 "Nuclear events in ethylene signaling: a transcriptional cascade
 mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1".
 Genes Dev. 12:3703-3714 (1998).
 RT
 CC -!- FUNCTION: Probable transcription factor that may be involved in
 the ethylene response pathway.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the EIN3 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF004215; AAC49748.1; -;
 DR EMBL; AC012679; AAG52067.1; -;
 DR EMBL; AY070044; AAL49801.1; -;
 DR EMBL; AY133839; AAM91773.1; -;
 DR EIR; E96764; E96764.
 DR TRANSFAC; T02652; -;
 DR InterPro; IPR006957; EIN3.
 DR Pfam; PF04873; EIN3; 1.
 KW Transcription regulation; Nuclear protein; DNA-binding; Coiled coil.
 FT DOMAIN 24 44 COILED COIL (POTENTIAL).
 SQ SEQUENCE 567 AA; 64041 MW; 308AFE4B3109594 CRC64;
 Query Match 61.9%; Score 39; DB 1; Length 567;
 Best Local Similarity 83.3%; Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 PRWPT 7
 Db 200 PPWPT 205
 RESULT 14
 ID YOCABEL STANDARD; PRT; 175 AA.
 AC Q23280;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 19.4 kDa protein ZC395.10 in chromosome III.
 GN ZC395.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]_SIMILARITY: Belongs to family 25 of glycosyl hydrolases.
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RL Connell M.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the p23 / wos2 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U13642; AAG00038.1; -;
 DR PIR; T27543; T27543.
 DR WormPep; ZC395.10; CE01436.
 DR InterPro; IPR008978; HSP20_chap.
 KW Hypothetical protein.
 FT DOMAIN 145 175 ASP/GLU-RICH.
 FT DOMAIN 165 168 POLY-GLU.
 SQ SEQUENCE 175 AA; 19431 MW; D5C136F30446E37A CRC64;
 Query Match 60.3%; Score 38; DB 1; Length 175;
 Best Local Similarity 55.6%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SPRWPTCL 9
 Db 81 TPWWPRL 89
 RESULT 15
 ID LYCH CHASP STANDARD; PRT; 211 AA.
 AC P00721;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE N.O.-diacetylmuramidase (EC 3.2.1.-) (Lysozyme CH).
 OS Chalaropsis sp.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 OC Helotiales; mitosporic Helotiales; Chalara.
 OX NCBI_TaxID=36534;
 RN [1]_SEQUENCE
 RX MEDLINE=75151523; PubMed=1168638;
 RA Felch J.W., Inagami T., Hash J.H.;
 RT "The N, O-diacetylmuramidase of Chalaropsis species. V. The complete
 amino acid sequence".
 RL J. Biol. Chem. 250:3713-3720 (1975).
 RL [2]
 RP ACTIVE SITE.
 RX MEDLINE=79005662; PubMed=567645;
 RA Fouché P.B., Hash J.H.;
 RT "The N,O-diacetylmuramidase of Chalaropsis species. Identification of
 aspartyl and glutamyl residues in the active site".
 RL J. Biol. Chem. 253:6787-6793 (1978).
 CC -!- FUNCTION: This enzyme has both lysozyme (acetylmuramidase) and
 diacetylmuramidase activities.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 heteropolymers of the prokaryotes cell walls.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- SIMILARITY: Belongs to family 25 of glycosyl hydrolases.
 DR PIR; A00876; MUKAD.
 DR InterPro; IPR008270; Glyco_hydro_25_AS.
 DR InterPro; IPR002053; Glyco_hydro_25.
 DR Pfam; PF01183; Glyco_hydro_25; 1.
 DR ProDom; PD004620; Glyco_hydro_25; 1.
 DR SMART; SM00641; Glyco_25; 1.
 DR PROSITE; PS00953; GLYCOSYL_HYDROL_F25; 1.
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme.
 FT ACT SITE 6
 FT ACT SITE 100 100 BY SIMILARITY.
 FT DISULFID 108 147
 SQ SEQUENCE 211 AA; 22413 MW; 379D758A393EC38C CRC64;
 Query Match 60.3%; Score 38; DB 1; Length 211;
 Best Local Similarity 50.0%; Pred. No. 37;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SPRWPTC 8

Db :| || :|

140 NPSWWSSC 147

Search completed: September 13, 2004, 14:19:41
Job time : 11.5263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2004, 14:08:15 ; Search time 8 Seconds
(without alignments)

354.958 Million cell updates/sec

Title: US-10-019-219A-2

Perfect score: 63

Sequence: 1 SPRWWTCL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	73.0	806	10 Q9ZW71	Q9zw71 arabidopsis
2	45	71.4	717	10 Q9XF33	Q9xf33 oryza sativ
3	43	68.3	122	16 Q8PG94	Q8pg94 xanthomonas
4	43	68.3	136	10 Q940S8	Q940s8 rosa hybrid
5	43	68.3	139	2 Q9EX86	Q9ex86 planobispor
6	43	68.3	145	12 Q9W895	Q9w895 hepatitis c
7	43	68.3	145	12 Q72124	Q72124 hepatitis c
8	43	68.3	145	12 Q72125	Q72125 hepatitis c
9	43	68.3	145	12 Q72123	Q72123 hepatitis c
10	43	68.3	145	12 Q92977	Q92977 hepatitis c
11	43	68.3	145	12 Q72126	Q72126 hepatitis c
12	43	68.3	237	4 Q9H7W0	Q9h7w0 homo sapien
13	43	68.3	300	10 Q84QD2	Q84qd2 nicotiana t
14	43	68.3	371	16 Q8PGK4	Q8pgk4 xanthomonas
15	43	68.3	383	10 Q9FWG2	Q9fwg2 oryza sativ
16	43	68.3	458	10 Q9FRJ1	Q9frj1 oryza sativ

17	43	68.3	458	10 Q7XCM0	Q7xcm0 oryza sativ
18	43	68.3	603	10 Q84QD1	Q84qd1 nicotiana t
19	43	68.3	615	10 Q948P3	Q948p3 cucumis mel
20	43	68.3	3010	12 Q92969	Q92969 hepatitis c
21	43	68.3	3010	12 Q92970	Q92970 hepatitis c
22	43	68.3	3010	12 Q92972	Q92972 hepatitis c
23	43	68.3	3010	12 Q02828	Q02828 hepatitis c
24	43	68.3	3010	12 Q92971	Q92971 hepatitis c
25	42	66.7	136	5 Q25300	Q25300 leishmania
26	42	66.7	151	16 Q8P5W0	Q8p5w0 xanthomonas
27	42	66.7	182	16 Q7V401	Q7v401 prochloroco
28	42	66.7	221	16 Q9JZ46	Q9jz46 neisseria m
29	42	66.7	221	16 Q9JU40	Q9ju40 neisseria m
30	42	66.7	377	10 Q9XEP9	Q9xep9 sorghum bic
31	42	66.7	387	16 Q7WKP0	Q7wkp0 bordetella
32	42	66.7	387	16 Q7W7A3	Q7w7a3 bordetella
33	42	66.7	577	5 Q96756	Q96756 dugesia tig
34	42	66.7	651	2 Q8KNF9	Q8knf9 micromonosp
35	42	66.7	858	5 Q27681	Q27681 leishmania
36	42	66.7	950	11 Q80TH1	Q80th1 mus musculu
37	41	65.1	103	12 Q9Q2P0	Q9q2p0 human herpe
38	41	65.1	122	10 Q8GUT2	Q8gut2 arabidopsis
39	41	65.1	123	17 Q9YDA5	Q9yda5 aeropyrum p
40	41	65.1	152	16 Q9L274	Q9l274 streptomyce
41	41	65.1	152	16 Q827V2	Q827v2 streptomyce
42	41	65.1	201	16 Q8PMR0	Q8pmr0 xanthomonas
43	41	65.1	207	5 Q8STY1	Q8sty1 encephalito
44	41	65.1	224	16 Q93S03	Q93s03 streptomyce
45	41	65.1	275	16 Q8FV55	Q8fv55 brucella su

ALIGNMENTS

RESULT 1

Q9ZW71 ID Q9ZW71 PRELIMINARY; PRT; 806 AA.
AC Q9ZW71;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE At2g43240 protein.
GN At2g43240.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
CX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004450; AAC64313.1; -.
DR PIR; G84863; G84863.
DR InterPro; IPR007271; Nuc_sug_transp.
DR Pfam; PF04142; Nuc_sug_transp; 1.
SQ SEQUENCE 806 AA; -89071 MW; 1F4D5ED6CBFDD89A CRC64;

```

Query Match          73.0%; Score 46; DB 10; Length 806;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWTCL 9
Db 676 PKWWTSL 683

RESULT 2
Q9XF33
ID Q9XF33 PRELIMINARY; PRT; 717 AA.
AC Q9XF33;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Oryza sativa (indica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39946;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Tetig;
RA Llaca V., Lou A., Young S., Messing J.;
RT "Microcollinearity in cereal genomes.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF128457; AAD27632.1; -.
DR Gramene; Q9XF33; -.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR00210; BTB POZ.
DR InterPro; IPR008974; Traf_dom.
DR Pfam; PF00651; BTB; 1.
DR SMART; SM00225; BTB; 1.
DR PROSITE; PS00097; BTB; 1.
KW Hypothetical protein.
SQ SEQUENCE 717 AA; 79014 MW; 90DBB88B78119F60 CRC64;

Query Match          71.4%; Score 45; DB 10; Length 717;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWWTCL 8
Db 302 PKWPHFC 308

RESULT 3
Q8PG94
ID Q8PG94 PRELIMINARY; PRT; 122 AA.
AC Q8PG94;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC3723.
GN XAC3723.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.F., Lemos E.G.M., Lemos M.V.F.,

```

```

RA Locali B.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.R.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463 (2002).
DR EMBL; AB012022; AAM38566.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 122 AA; 14026 MW; AF73F08878C62E20 CRC64;

Query Match          68.3%; Score 43; DB 16; Length 122;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWWPTC 8
Db 4 RWWPAC 9

RESULT 4
Q940S8
ID Q940S8 PRELIMINARY; PRT; 136 AA.
AC Q940S8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE EIN3-like transcription factor (Fragment).
GN EIN3.
OS Rosa hybrid cultivar.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
OX NCBI_TaxID=128735;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller R., Owen C.A., Stummann B.M.;
RT "partial sequence of Rosa hybrida cultivar mRNA for EIN3-like
RT transcription factor.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052825; AAL14267.1; -.
DR InterPro; IPR006957; EIN3.
DR Pfam; PF04873; EIN3; 1.
DR NON TER 1
FT NON TER 136
SQ SEQUENCE 136 AA; 15267 MW; B77FAD5FBE1383B9 CRC64;

Query Match          68.3%; Score 43; DB 10; Length 136;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPRWWPT 7
Db 106 SPWWPT 112

RESULT 5
Q9EX86
ID Q9EX86 PRELIMINARY; PRT; 139 AA.
AC Q9EX86;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative peptide synthetase (Fragment).
OS Planobispora rosea.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Streptosporangaceae; Planobispora.
OX NCBI_TaxID=35762;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=ATCC 53733;
 RX MEDLINE=20535709; PubMed=11085259;
 RT Sosio M., Bossi E., Bianchi A., Donadio S.;
 RA "Multiple peptide synthetase gene clusters in actinomycetes.";
 RL Mol. Gen. Genet. 264:213-221(2000).
 DR EMBL; AJ276363; CAC01622.1; -;
 DR HSSP; P14687; IAMU.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR00873; AMP-bind.
 DR InterPro; IPR006163; PP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; PP-binding; 1.
 DR PROSITE; PSS0075; ACP_DOMAIN; 1.
 FT NON TER 1
 FT NON TER 139
 SQ SEQUENCE 139 AA; 15029 MW; 1F2489785FD715C6 CRC64;

Query Match 68.3%; Score 43; DB 2; Length 139;
 Best Local Similarity 85.7%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPWWPTCL 7
 |||:|
 Db 31 SPGWPT 37

RESULT 6
 Q9WB95 ID Q9WB95 PRELIMINARY; PRT; 145 AA.
 AC Q9WB95;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 NCBI_TaxID=11103;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b
 RT are infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054267; AAC15740.1; -;
 DR EMBL; AF054263; AAC15736.1; -;
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR Pfam; PF00998; HCV RdRP.
 DR InterPro; IPR002166; HCV RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KW Transferase.
 FT NON TER 1
 FT NON TER 145
 SQ SEQUENCE 145 AA; 16175 MW; 8157D280215C2252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWWPTCL 9
 |||:|
 Db 123 PRWPLCL 130

RESULT 7
 O72124 ID O72124 PRELIMINARY; PRT; 145 AA.
 AC O72124;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 NCBI_TaxID=11103;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054264; AAC15737.1; -;
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR002166; HCV RdRP.
 DR Pfam; PF00998; Viral RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KW Transferase.
 FT NON TER 1
 FT NON TER 145
 SQ SEQUENCE 145 AA; 16087 MW; 83AAFD4C3C9DE252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWWPTCL 9
 |||:|
 Db 123 PRWPLCL 130

RESULT 8
 O72125 ID O72125 PRELIMINARY; PRT; 145 AA.
 AC O72125;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 NCBI_TaxID=11103;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054266; AAC15739.1; -;
 DR GO; GO:0019012; C:virion; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0006350; P:transcription; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR002166; HCV RdRP.
 DR Pfam; PF00998; Viral RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KW Transferase.
 FT NON_TER 1
 FT NON_TER 145
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16175 MW; 8157C2CD7999E252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPTCL 9
 |||:|
 Db 123 PRWPTCL 130

RESULT 9
 ID 072123 PRELIMINARY; PRT; 145 AA.
 AC 072123;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054262; AAC15735.1; -.
 DR GO; GO:0019012; C:viral genome replication; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003668; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR002166; HCV RdRP.
 DR Pfam; PF00998; Viral RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KW Transferase.
 FT NON_TER 1
 FT NON_TER 145
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16145 MW; 8157D29C3C9DE252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPTCL 9
 |||:|
 Db 123 PRWPTCL 130

RESULT 10
 ID 092977 PRELIMINARY; PRT; 145 AA.
 AC 092977; 092978;
 DT 01-NOV-1998 (TRENBLrel. 08, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b
 RT are infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054260; AAC15733.1; -.
 DR EMBL; AF054261; AAC15734.1; -.
 DR EMBL; AF054265; AAC15738.1; -.
 DR GO; GO:0019012; C:viral genome replication; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003668; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR002166; HCV RdRP.
 DR Pfam; PF00998; Viral RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KW Transferase.
 FT NON_TER 1
 FT NON_TER 145
 FT NON_TER 145
 SQ SEQUENCE 145 AA; 16235 MW; 8157D290205C2252 CRC64;

Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PRWPTCL 9
 |||:|
 Db 123 PRWPTCL 130

RESULT 11
 ID 072126 PRELIMINARY; PRT; 145 AA.
 AC 072126;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Genome polyprotein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788;
 RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo.";
 RL Virology 244:161-172(1998).
 DR EMBL; AF054268; AAC15741.1; -.
 DR GO; GO:0019012; C:viral genome replication; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003668; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR002166; HCV RdRP.
 DR Pfam; PF00998; Viral RdRP; 1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;

KW Transferase. 1 1
 FT NON_TER 145 145
 SQ SEQUENCE 145 AA; 16175 MW; 8157C2CD7999E252 CRC64;
 Query Match 68.3%; Score 43; DB 12; Length 145;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PRWPTCL 9
 Db 123 PRWPTCL 130

RESULT 12
 Q9H7W0 ID Q9H7W0 PRELIMINARY; PRT; 237 AA.
 AC Q9H7W0
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Hypothetical protein FJ114202.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX da Silva A.C.R., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Nanmoto K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024264; BAB14864.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 237 AA; 25615 MW; 2CE39E9D320B3863 CRC64;

Query Match 68.3%; Score 43; DB 4; Length 237;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWPCT 8
 Db 99 WWPCT 103

RESULT 13
 Q84QD2 ID Q84QD2 PRELIMINARY; PRT; 300 AA.
 AC Q84QD2
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE EIL3.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Petit Havana SRI;
 RA Rieu I., Mariani C., Weterings K.;
 RT "Tobacco EIN3-like cDNAs."
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY248905; AAP03999.1; -.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR006957; EIN3.
 DR Pfam; PF04873; EIN3; 1.

DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 SQ SEQUENCE 300 AA; 34279 MW; 9CC7A44EFB81442B CRC64;
 Query Match 68.3%; Score 43; DB 10; Length 300;
 Best Local Similarity 85.7%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPRWPT 7
 Db 212 SPWPT 218

RESULT 14
 Q8PGK4 ID Q8PGK4 PRELIMINARY; PRT; 371 AA.
 AC Q8PGK4
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Peptidase.
 GN XAC3612.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., Amaranal A.M., Bertolini M.C., Camargo L.E.A., P.,
 RA Amarante G., Cannan F., Cardoso J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 DR EMBL; AE012011; AM38455.1; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR003137; PA.
 DR Pfam; PF02225; PA; 1.
 KW Complete proteome.
 SQ SEQUENCE 371 AA; 41061 MW; 1384E69BC0FABC96 CRC64;

Query Match 68.3%; Score 43; DB 16; Length 371;
 Best Local Similarity 83.3%; Pred. No. 74;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RWPTC 8
 Db 2 RWPAC 7

RESULT 15
 Q9FWG2 ID Q9FWG2 PRELIMINARY; PRT; 383 AA.
 AC Q9FWG2
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OSUNB001511.26.

OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=sv. Nipponbare;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zisemann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNB0015111 genomic sequence.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC EMBL; AC051633; AG13599.1; -.
 DR HSP; F35813; LA60.
 DR Gramine; Q9FWG2; -.
 DR GO; GO:0008287; C:protein serine/threonine phosphatase complex; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0000287; F:magnesium ion binding; IEA.
 DR GO; GO:0004722; F:protein serine/threonine phosphatase activity; IEA.
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR000222; PP2C.
 DR InterPro; IPR001932; PP2C-like.
 DR Pfam; PF00481; PP2C; 1.
 DR SMART; SM00332; PP2C; 1.
 DR SMART; SM00331; PP2C-SIG; 1.
 DR PROSITE; PS01032; PP2C; 1.
 KW Hypothetical protein; Hydrolase; Magnesium.
 SQ SEQUENCE 383 AA; 41899 MW; 9B58D7BCD5217B81 CRC64;

Query Match 68.3%; Score 43; DB 10; Length 383;
 Best Local Similarity 71.4%; Pred. NO. 76;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RWWPTCL 9
 Db 14 RWWPICV 20

Search completed: September 13, 2004, 14:22:16
 Job time : 10 secs